

AD-A280 360



AEOSR-TR- 94 0350

19 NOV 93

1

THE ROLE OF PORTABLE INSTRUMENTATION IN MONITORING A COMPREHENSIVE TEST BAN TREATY

Brian W. Stump*
Department of Geological Sciences
Southern Methodist University
Dallas, Texas 75275-0395

Approved for public release;
distribution unlimited.

Florence Rivière-Barbier
SAIC/ Center for Seismic Studies
Arlington, Virginia

Igor Chernoby
Experimental Methodological Expedition
Institute of Physics of the Earth
Obninsk, Russia

Karl Koch
Department of Geological Sciences
Southern Methodist University
Dallas, Texas 75275-0395

Seismic monitoring of a Comprehensive Test Ban Treaty (CTBT) may require the detection, location and identification of seismic events as small as $m_b = 2.5$ (Wallace *et al.*, 1992) in limited areas of interest. With the emphasis placed on such an agreement by the current Administration, it is important to assess the complexity of the proposed task. The smallest events that must be discriminated from nuclear explosions include those associated with human activities such as construction and mining. These small magnitude events may be recorded by only a few regional stations (OTA Report, 1988). The lowest magnitude level to which monitoring must be accomplished is dependent on the quantification of various evasion scenarios, the most important of which may be decoupling (Murphy *et al.*, 1993; Stevens *et al.*, 1991).

To quantify the size of the monitoring problems, one must first relate the explosive yield of mining explosions to a magnitude measure. Israelson and Carter (1991) compare total explosive weight in ripple-fired explosions to M_L and suggest that in Fennoscandia a 25-50 ton explosion would have a M_L of 2.5 with a coupling scatter as great as a factor of 6-8. The magnitude-yield curves reported by Stevens *et al.* (1991) for unsaturated and saturated geologic materials at NTS predict m_b 's for a contained 25 ton nuclear explosion of 2.04 to 2.64. Reamer and Stump (1991) compared near-source and regional measurements of a series of surface chemical explosions in the Western US. The 150 ton explosion in the series, assigned a M_L of 3.1 in the *Preliminary Determination of Epicenters* by the USGS, is consistent with these other results. These observations and models suggest that a monitoring threshold of 25-50 tons for ripple-fired explosions would be consistent with a magnitude threshold near 2.5. The number of man made events greater than 50 tons in the US is 10,000 (Richards *et al.*, 1992) with one shot per day over 200 tons.

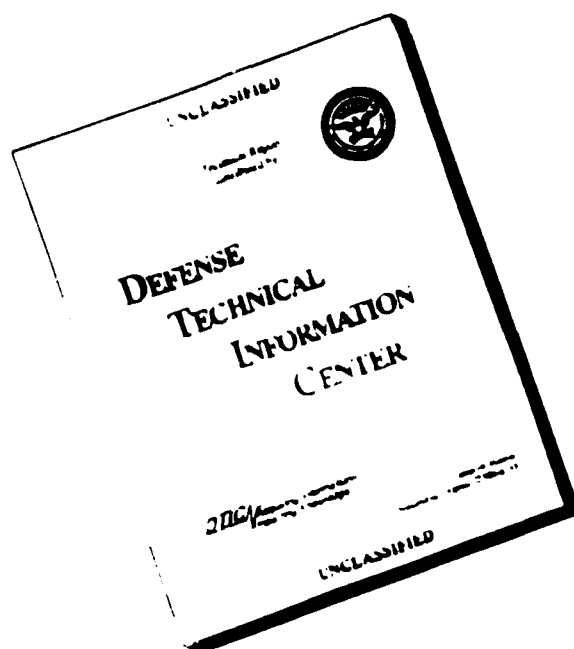
94-18484

26092

94 6 14 148

A-1

DISCLAIMER NOTICE



THIS REPORT IS INCOMPLETE BUT IS THE BEST AVAILABLE COPY FURNISHED TO THE CENTER. THERE ARE MULTIPLE MISSING PAGES. ALL ATTEMPTS TO DATE TO OBTAIN THE MISSING PAGES HAVE BEEN UNSUCCESSFUL.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE JUN 15 1994			5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 94 0350		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) SMU #5-25155			7a. NAME OF MONITORING ORGANIZATION		
6a. NAME OF PERFORMING ORGANIZATION Southern Methodist University Dept. of Geological Sciences		6b. OFFICE SYMBOL (If applicable)		7b. ADDRESS (City, State, and ZIP Code)	
6c. ADDRESS (City, State, and ZIP Code) Southern Methodist University Dept. of Geological Sciences Dallas, TX 75275-0395		8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR/NL Stan Dickenson		8b. OFFICE SYMBOL (If applicable)	
8c. ADDRESS (City, State, and ZIP Code) AFOSR/NL 110 Duncan Avenue Suite B115 Bolling AFB DC 20332-0001		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER			
11. TITLE (Include Security Classification) The Role of Near Source Phenomenology on Regional Observations		10. SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO. 61102F		PROJECT NO. 2309	
		TASK NO. AS		WORK UNIT ACCESSION NO.	
12. PERSONAL AUTHOR(S) B. Stump, D. Anderson, G. Min, K. Koch, F. Riviere-Barbier, I. Chernoby					
13a. TYPE OF REPORT Annual		13b. TIME COVERED FROM 01/93 TO 02/94		14. DATE OF REPORT (Year, Month, Day) 940413	
15. PAGE COUNT 265					
16. SUPPLEMENTARY NOTATION AFOSR #F49620-93-1-0146					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Mining explosion, nuclear explosion, seismic source function		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)					
<p>This report documents two efforts undertaken during the past 12 months. The first describes a combined near-source/regional monitoring of a series of mining blasts in Southern Russia. The second contribution describes a theoretical investigation of nuclear explosion source model resolution using near-source seismic data in a nonlinear inversion scheme.</p>					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION		
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Eugene Herrin			22b. TELEPHONE (Include Area Code) 214/768-2744		22c. OFFICE SYMBOL

8 MAY 1994

The discrimination between earthquakes, chemical mining explosions and nuclear explosions using regional seismic waves ($P/S/L_g$ ratios, spectral scalloping, frequency content) has been shown to be strongly region dependent (Patton, 1993; Baumgardt and Der, 1993). The establishment of a physical framework for discriminants is important if successful techniques developed in one region are to be reliably transported and used in another location. Quick acquisition of region specific data, such as information related to crust and upper mantle velocity model, wave propagation characteristics and mining practices of interest, is required for practical implementation of a monitoring system. The utilization of portable instrumentation provides the opportunity to acquire such information in the direct vicinity of the source as well as at regional distances. Digital data acquisition systems developed under the PASSCAL program linked with GPS clocks provide the necessary equipment for integrated near-source and regional studies.

An experiment was executed during the last two weeks of August 1993 to test the applicability of such a seismic monitoring system combining near-source and regional data. It was conducted in and around an ore mine located in Southern Russia at Tyrnyauz in the Caucasus Mountains (Cover Photo and Figure 1). The goals of the deployment were: (1) document blasting practices; (2) quantify the coupling of seismic energy at close-in distances; and (3) resolve regional propagation path effects. The experimental work involved contributions from three institutions in Russia: *Experimental Methodological Expedition (Obrinsk)*, *Institute for Dynamics of the Geospheres (Moscow)*, *Institute of Physics of the Earth (Moscow)*; and two institutions in the United States: *Southern Methodist University (Dallas, TX)* and *Center for Seismic Studies (Arlington, VA)*. Multiple types of observations were made of the explosions and included near-source and regional seismic ground motions, high speed film and video, electromagnetic measurements and field documentation. These data provided additional constraints to the seismic source and were used to interpret the adequacy of discriminants often applied only to regional seismograms. The focus of this note will be on the seismic observations, the field documentation and the video records from the explosions.

Validation of mining and blasting practices through direct field observations is identified as 'ground truthing.' These direct observations are compared to official records of blasting practices maintained by the mine. The types of information labeled as blasting practices include the size and types of boreholes, amount and type of explosive, and method and timing of detonation.

TYRNYAUZ MINE AND MINING PRACTICES

The Tyrnyauz mine is located in the Kabardino-Balkaria Republic of Russia close to the Georgian border (Figure 1). This particular mine was chosen for study because of a history of large explosions, two high-quality regional arrays, the occurrence of near-by earthquakes and cooperation with the mine operators. The city of Tyrnyauz has a population of 10,000 with half of these people employed in either the mining or the processing activities. Mineral exploitation began in 1940 in both underground and near-surface (cover photo) mines between 2400 and 3000 m. In the underground operation, over 150 km of 5.5 m diameter tunnels have been excavated. Both tungsten and molybdenum are extracted from the various metamorphic rocks present in this part of the Caucasus. The purpose of the blasting is to fragment the rock to sizes of 900 mm or less. These rock fragments are further reduced in size to 100-350 mm when they are dropped down a 700 m deep well for processing at lower elevations in the mine.

Typically both near-surface and underground production shots are detonated on Sunday mornings. The smaller underground explosions are completed first and consist of one to

several charges detonated simultaneously. A near-surface explosion can involve many separate borehole explosions in rows on multiple benches or at different elevations. The individual shots within each row are detonated simultaneously with 20 to 40 ms delays between rows depending on the borehole depths. Boreholes are partially filled with a granular explosive consisting of 71% ANFO covered with an aluminum powder. The detonation is initiated with an electronic blasting machine which in turn ignites detonating cord with a burn rate of 7000 m/s. The purpose of the explosives is to fragment the rock with little or no concern for mass movement. As a result of this philosophy, the blasts tend to bulk the material moving it primarily in the vertical direction. Engineering records at the mine for 11 August 1991 to 28 August 1993 indicate that 6 surface explosions had yields in excess of 50 tons with an average explosion size of 33 tons for this time period. Underground and near-surface explosions were observed on 22 and 29 August 1993. On both days the underground explosions were detonated first with the near-surface following approximately one hour later. The sizes of these explosions were relatively small: 18.9 and 5.8 tons for the underground explosions and 25.3 and 7.3 tons for the surface explosions. The underground explosions consisted of one (29 Aug) and four (22 Aug) individual charges detonated simultaneously.

Official design records for the near-surface blasts were obtained from the mine engineers. Comparison between these records and the actual field deployment of explosives as well as video and photographic documentation of the near-surface explosion revealed wide discrepancies between the documented and actual explosions. Figure 2 compares the planned near-surface blast for 22 August according to official mine records (blue and green symbols) with that detonated as determined by field documentation (blue and white symbols). The total number of boreholes in the actual blast was reduced from that planned as well as the amount of explosive per hole. In addition, thirty bags of explosive (white spheres in Figure 2) were added to the near-surface explosion by draping them across large surface rocks. These bags (42 kg of explosive each) were not placed in boreholes and were intended to fracture large boulders remaining from previous blasts. The time between the rows of boreholes was increased from a planned delay of 25 ms to 40 ms. These changes resulted in a reduction in total explosive charge from 43.3 tons in the official records to an actual yield of 25.3 tons. A significant air blast was introduced from the bags of explosives placed on the boulders and the lack of stemming in each emplacement hole. The discrepancy between official mine records and actual blasting practice illustrates the importance of near-source monitoring of mining practices in order to fully assess source effects on regional seismograms. This 'ground truthing' provides the quantitative information that can be used to separate source and propagation path effects unambiguously at regional distances. Reliance upon official mining records may be misleading if this experience is typical of other mines. The changes that were introduced were brought about by the availability of explosive resources on the day of the shot and the local site geology as interpreted by the blaster. It is not unreasonable to expect similar variations in other mining operations.

Another aspect of the field documentation was the utilization of video and high speed film to determine the timing and regularity of the explosions. Figure 3 displays four video frames (sampled at 16.67 ms/frame) of the near-surface explosion on 22 August. As indicated in Figure 2, this blast occurred on two levels or benches. The first frame illustrates blasts on the first bench. The boreholes are not back-filled to the surface so that explosive by-products can be readily identified in the images. The second frame captures the detonation of bags of explosives on the first bench. These explosions are indicated by the bright orange images. The third frame illustrates the initiation of the first row on the second bench although all the boreholes do not fire simultaneously, probably as a result of variations in the individual blasting caps in each hole. A number of authors have suggested that regular delay times between individual charges or rows of charges in this case may lead

to consistent spectral scalloping in the Fourier spectra of the seismograms (Hedlin, Minster and Orcutt, 1989). These photos indicate that there may be variation between the design and actual shot times thus randomizing the spectral characterization and possibly degrading this discriminant.

SEISMIC INSTRUMENTATION

Near-instantaneous monitoring of man-made seismic sources requires a set of rugged and easily deployed instruments with relatively wide dynamic range. In addition, the data recovered from such a system must be combined in a timely manner with existing permanent regional seismic networks. These experimental goals led to the assembly of a portable instrumentation system for the near-source observations based upon two, six-channel Refraction Technology data acquisition systems (DAS), model 72-06. In order to span the range of ground motions expected in the near-source region, two sets of sensors were deployed with each DAS and included a three-component set of Terra Technology accelerometers and three-component Sprengnether S-6000 2 Hz seismometers. Timing and location information for each instrument was provided by a GPS receiver, making the near-source data available for immediate integration with the regional data. Sixteen-bit data were recovered at 500 samples/s in order to characterize the near-source ground motions. This data provided a wide-band picture of the source that could be compared to the other near-source observations such as the high speed photography.

Regional seismic data were recorded by Experimental Methodological Expedition (EME) operated facilities: two regional telemetry networks (RSS, installed by EME and a Nanometrics telemetry system installed by Lamont); the Kislovodsk micro-array (installed by CSS); and the broadband IRIS/IDA seismic station (installed by UCSD) (Figure 1). The RSS network includes 7 stations equipped with CM3-KB three-component seismometers and a data acquisition and recording system (designed by EME) with a sampling rate of 128 samples/s. The system has flat velocity response between 0.4 and 20 Hz. The Lamont system consists of seismometers collocated with RSS instruments and characterized by a flat velocity response between 0.2 and 24 Hz with a sample rate of 60 samples/s. The Kislovodsk 4-element micro-array with an aperture of 300 m is equipped with Teledyne-Geotech GS-13 seismometers -- three-components at the middle point and vertical only at the periphery. The instrument response is flat in velocity from 0.5 to 10 Hz and the data are sampled at 40 samples/s. The IRIS/IDA seismic station has three-component STS-1 seismometers with a flat velocity response between 0.003 and 5 Hz and is sampled at 20 samples/s.

NEAR-SOURCE DATA

The near-source data provide the opportunity to evaluate time and frequency domain differences between the simultaneous underground explosions and the ripple-fired near-surface explosions. Figure 4 compares the 22 August vertical velocity records from the near-surface and underground explosions at one of the near-source stations (S2). A number of source characteristics are immediately evident. First, the increased low frequency content of the near-surface explosion signal relative to the underground can be observed in both the time and frequency domain. The near-surface explosion spectrum is larger by as much as an order of magnitude in the frequency band of 1 to 5 Hz. The spectra from the two explosions merge at the higher frequencies although there is still considerable variation between the two at a given frequency. The total duration of the surface explosion is close to 200 ms and would predict a spectral hole at 5 Hz from this temporal window and suggests that source duration controls the spectral character in the 1 to 5 Hz band. Spectral interference from the interaction of the waveforms generated by each row is harder to identify in the spectra and may reflect the scatter in the individual detonations as identified in the video records. As noted in the explosion discussion, a significant variation from US blasting practices was the inclusion of free surface explosions in the mining blast and the lack of stemming in the emplacement holes. The high-frequency, late-time arrival on the vertical component of the near-source data is evidence of

this air blast. Monitoring of such arrivals may be useful in identifying similar types of blasting practices.

REGIONAL DATA

The regional observations from the same explosions allow one to directly assess the effect of propagation path on the source signatures identified in the near-source data. Comparisons between the underground and near-surface explosions on 22 August are made at the regional stations KNG (28 km), KIV (65 km) and GUM (67 km) in Figure 5. The time series from the surface explosion at each of these regional stations are enriched in low frequency energy relative to the seismograms from the underground explosion. Inspection of the whole record spectra accompanying each waveform illustrates that the surface explosion is again enriched to about 5 Hz where the spectra from the two events merge. This comparison confirms that the increased energy from the near-surface blasts, identified in the near-source observations, is also reflected in the regional waveforms. These data suggest that bandwidth measures of regional signals may be used to separate different types of above and underground explosions. Such a discriminant would rely on relative wide band data, out to 10 Hz or beyond in this example.

The repeatability of the source excitation is important if pattern recognition is to be used to separate source types at regional distances. Comparison of the regional signals at three stations (GUM, KIV, KNG) from the underground explosions on 22 and 29 August (Figure 6) illustrates the strong similarity in bandwidth and arrivals from these two sources. Despite the known yield differences (18.9 and 5.8 tons) these records suggest that pattern recognition procedures as proposed by Rivière-Barbier and Grant (1993) might be successful in identifying events of a similar geometry. The differences identified in the near-surface and underground shot (Figure 4 and 5) argues that subtle changes in source depth and spatial or temporal characterization might also be identified with comparable techniques.

Regional arrival time data were used to locate the two explosions on 22 August in order to investigate location bias introduced by utilization of a regional 1D velocity model. The regional locations of the explosions are within 1 km of those determined by the field investigation (Figure 1). This comparison emphasizes the value of selected near-source observations for regional network calibration.

CONCLUSIONS/IMPLICATIONS

The detection, location and identification of small seismic events will increase in importance if a Comprehensive Test Ban Treaty is implemented. This experiment has illustrated the utility of combined near-source and regional observations in studying unusual or unidentified events. Digital data acquisition systems in combination with a GPS provide the means for a rapid deployment of portable instrumentation that can quickly be integrated with an existing permanent array. The availability of Internet services further provides for rapid access to the data following the experiment. Correlation and distribution of both the regional and near-source data were performed from KIV the day of the explosions. Anomalous events identified by regional signals under a CTBT can be investigated with a system such as that deployed at Tyrnyauz. The near-source observations in combination with field documentation will provide additional data for improved event identification as a construction or mining activity. Studies such as this one can be used to identify important physical processes in the source region (total source duration and source depth in this case) that contribute to regional observations. The experiment has also identified significant variations between documented and actual blasting practices and suggests that care should

be applied when using formal blasting records from a mining operation in the interpretation of regional seismic records.

REFERENCES

- Baumgardt, D. R. and Z. Der, Investigation of regional seismic discriminants using visualization and statistical analysis methods in the intelligent seismic event identification system, in the Proceedings of the 15th Annual PL/AFOSR/ARPA Seismic Research Symposium, 8-10 September 1993, Vail, Colorado.
- Hedlin, M. A. H., J. B. Minster and J. A. Orcutt, The time-frequency characteristics of quarry blasts and calibration explosions recorded in Kazakhstan, USSR, *Geophys. J. Int.*, **99**, 109-121, 1989.
- Israelson, H. and J. Carter, Analysis of high frequency seismic data, PL-TR-91-2032, Phillips Laboratory, Hanscom AFB, MA, 1991.
- Murphy, J. R., J. Stevens and N. Rimer, Theoretical simulation analysis of seismic signals from decoupled explosions in spherical and ellipsoidal cavities, *EOS, Transactions, American Geophysical Union*, **74**, p 58, April 20, 1993.
- Patton, H. J., Discrimination of low magnitudes: Summary of potential and recent results, *EOS, Transactions, American Geophysical Union*, **74**, p 58, April 20, 1993.
- Reamer, S. K. and B. W. Stump, Source parameter estimation for large, bermed, surface chemical explosions, *Bull. Seism. Soc. Am.*, **82**, 406-421, 1992.
- Richards, P. G., D.A. Anderson, and D. W. Simpson, A survey of blasting activities in the United States, *Bull. Seism. Soc. Am.*, **82**, 1416-1433, 1992.
- Rivière-Barbier, F. and L. Grant, Identification and location of closely spaced mining events, *Bull. Seism. Soc. Am.*, **83**, 1527-1546, 1993.
- Stevens, J. L., J. R. Murphy and N. Rimer, Seismic source characteristics of cavity decoupled explosions in salt and tuff, *Bull. Seism. Soc. Am.*, **81**, 1272-1291, 1991.
- U.S. Congress, Office of Technology Assessment, *Seismic Verification of Nuclear Testing Treaties*, OTA-ISC-361 (Washington, DC: U.S. Government Printing Office, May 1988)
- Wallace, T., R. Blandford, A. Dainty, R. Lacoss, R. Maxion, A. Ryall, B. Stump and C. Thurber, *Report on the DARPA Seismic Identification Workshop, 18-19 May 1992*, 28 pp.

Acknowledgments: Funding for this experiment was provided by ARPA under contracts F29601-92-C-005 (CSS), F29601-91-D-DB20 (SMU), F19628-93-K-0016 and by AFOSR under grant F49620-93-1-0146. Special thanks go to the EME staff at Kislovodsk for their field support, to the operators of the Tyrnyauz mine (particularly German Theodorovich Kazan), and to David Anderson at SMU for the image processing.

*Brian Stump now at LANL, EES-3, MS-C335, Los Alamos, NM 87545. email stump@beta.lanl.gov. 505-6671004.

FIGURES

Cover Photo: Photo of the Tyrnyauz surface mine with snow capped Caucasus in the background. The explosions reported in this note were located to the far right of the photo.

Figure 1: Regional stations, mine, and blast locations are given in black. The local area around the Russian and Georgian border (purple) is illustrated. The error ellipse for the location of the explosion (gray) is based on the size of the estimated error in arrival time that was assumed to be 0.5 s for P phases and 1 s for S phases.

Figure 2: Three dimensional layout of the benches where the 22 August surface explosion was detonated with the design and actual explosion arrays displayed. Actual boreholes are represented by blue symbols, 42 kg bags of explosives placed on boulders at the surface by the white spheres, and planned but undetonated boreholes by the green symbols.

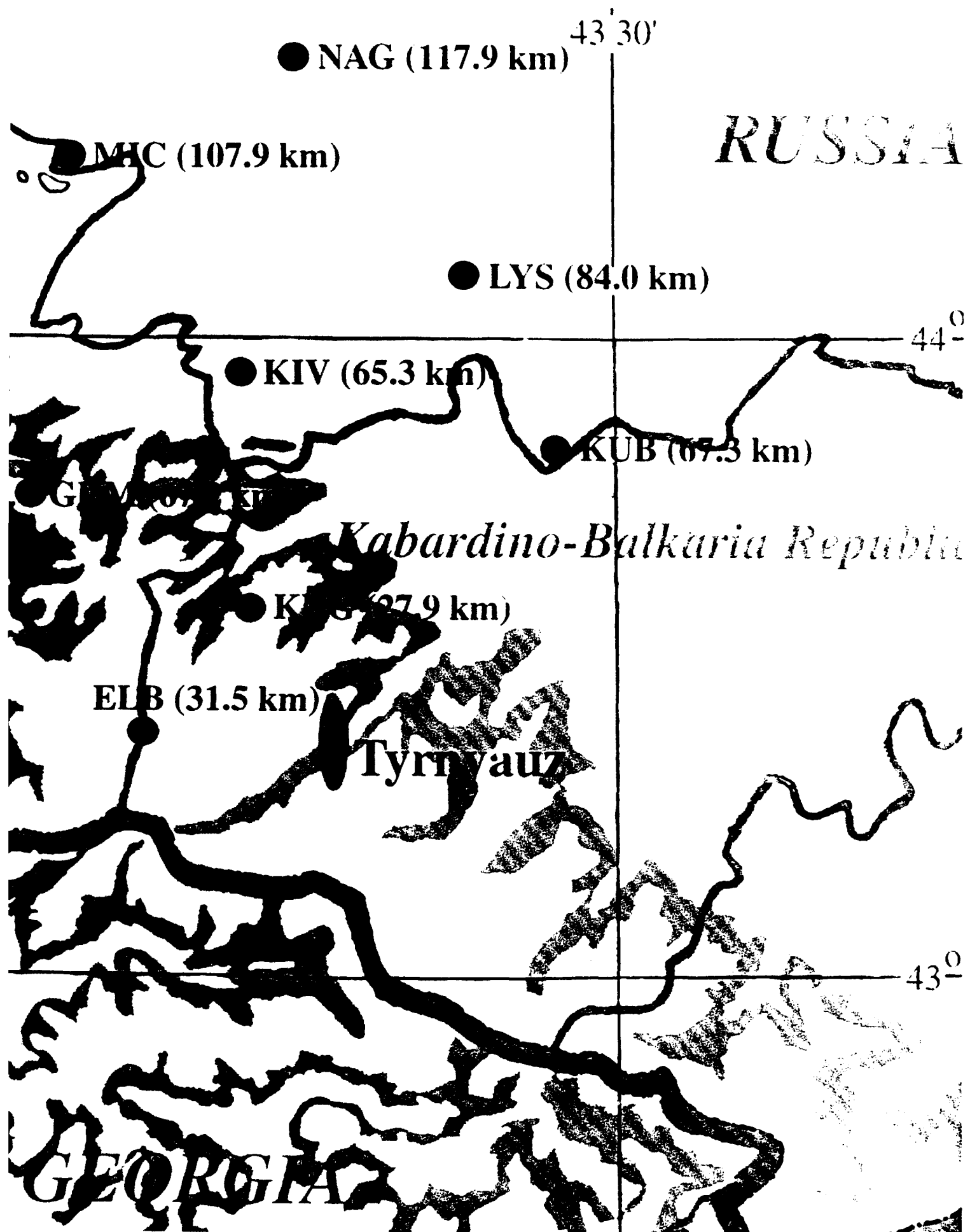
Figure 3: Composite image of 4 frames from the video characterizing the surface explosion on 22 Aug (time from the beginning of detonation is given in the upper left corner). These images display the first borehole detonations on the first bench (Frame 1), some of the bags detonated at the surface (Frame 2) and the detonations on the second bench (Frame 3 and 4).

Figure 4: Comparison of the vertical velocity records at one near-source station (S2) from the near-surface (25.3 tons) and underground (18.9 tons) explosions on 22 August. The corresponding whole record spectra are shown to the left.

Figure 5: Comparison of the near-surface and underground explosion (22 August) records at three of the regional seismic stations. The corresponding whole record spectra are shown to the right.

Figure 6: Comparison of the regional records (1 Hz high pass filtered) at three stations from the underground explosions on 22 and 29 August.





100 ms



166 ms



232 ms



315 ms

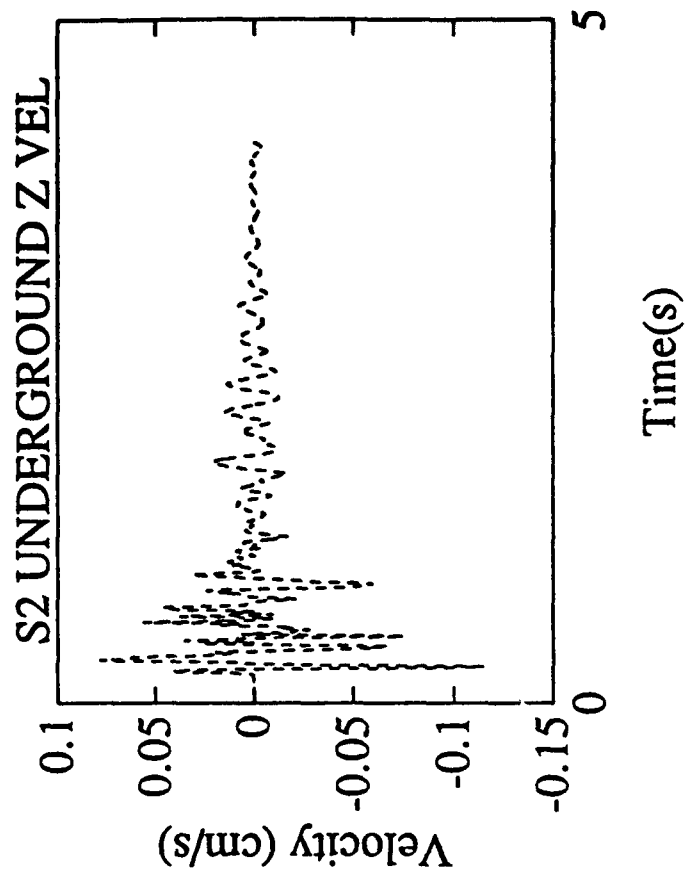
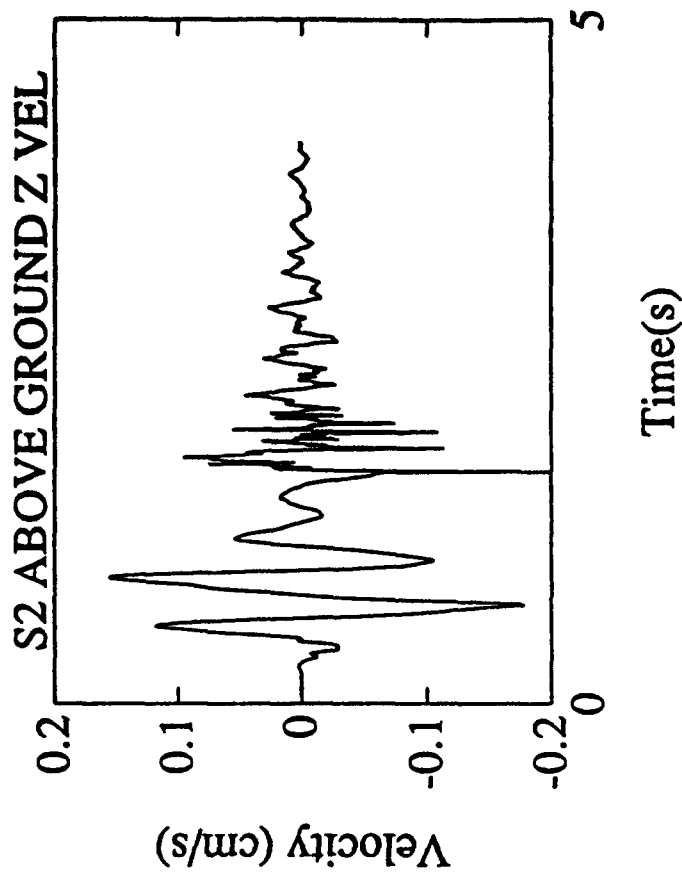
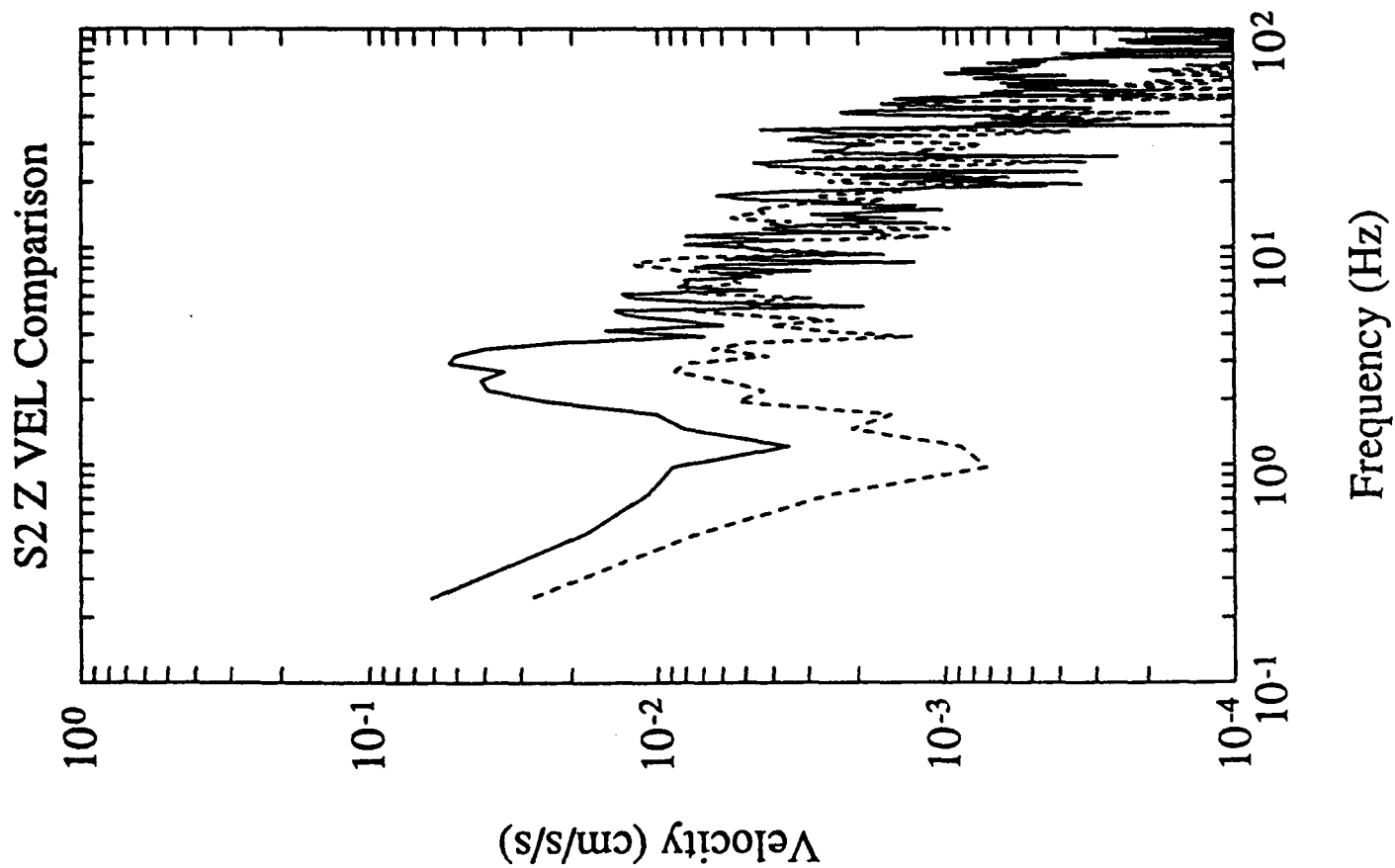


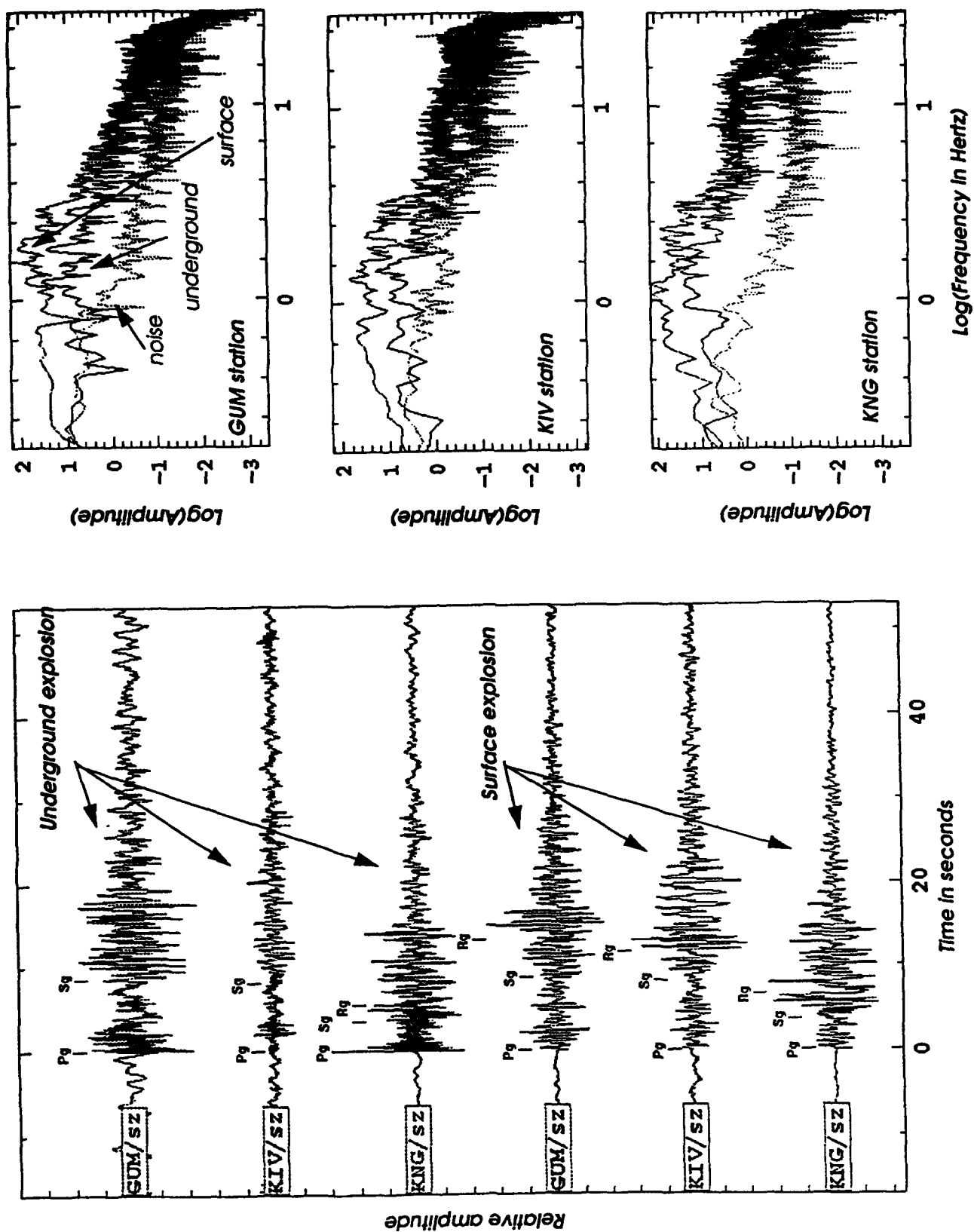


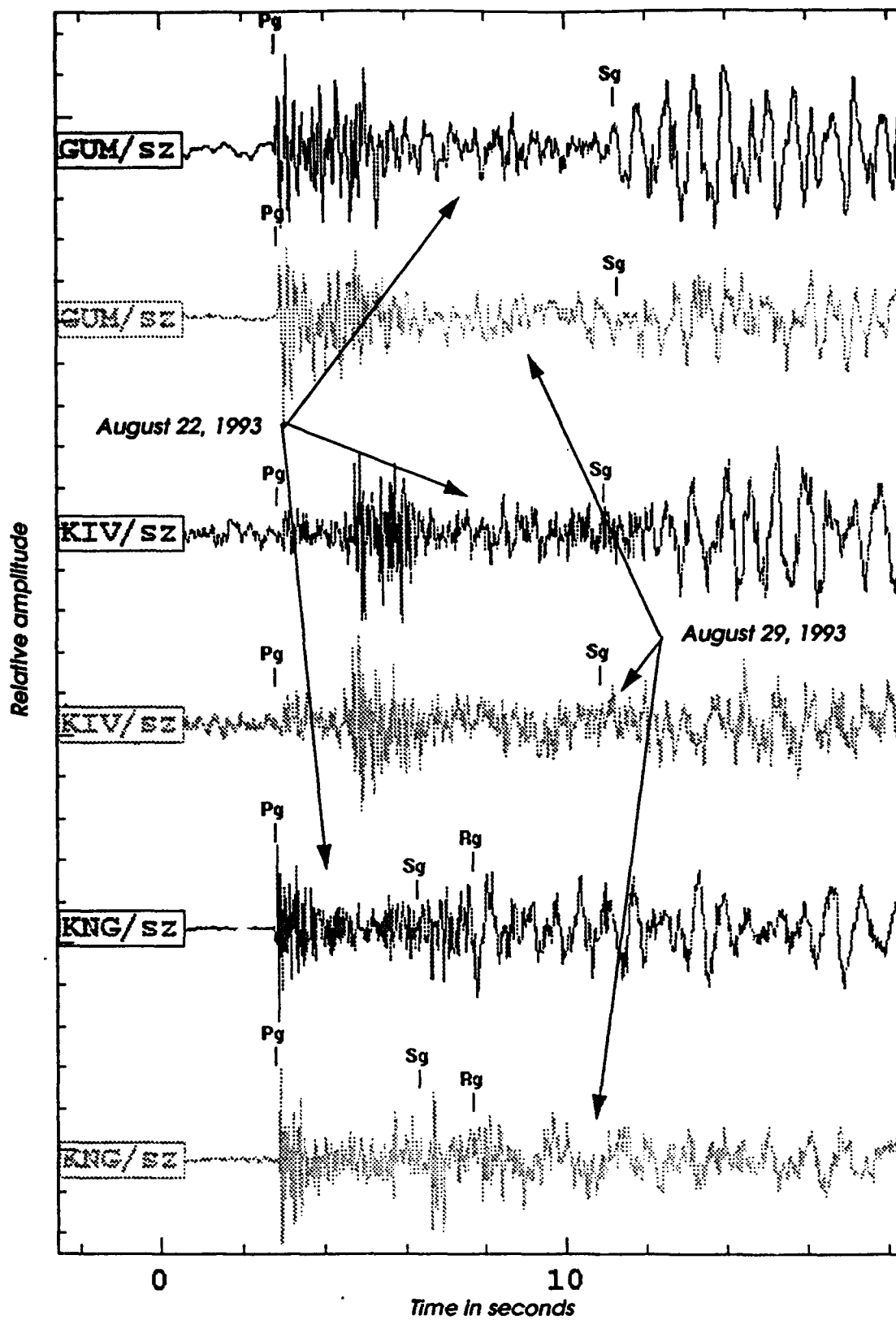
• 42 kg. 5'10"

• 294 kg. 5'10"

• 160 kg. 5'10"







CHAPTER 1

INTRODUCTION

The main objective of this research is to calculate seismic source parameters for nuclear explosions. The seismic source can be represented either mathematically by the moment tensor or physically by a seismic source function. Moment tensor, which is expressed by the 9 components of 3×3 matrix but generally has six unknowns because of the conservation of the angular momentum, will be a more general representation of the source including both earthquakes and explosions. In the special case of an explosion, the moment tensor can be reduced to a function of pressure with the assumption of isotropic normal stresses and no shear stress. This isotropic representation of seismic parameters is more concise and simpler than the full moment tensor expression.

Seismic source parameters, which will be discussed in a subsequent section, are related to the various physical parameters theoretically and empirically. For example in an explosion, yield(W), one of the physical parameters, can be expressed as a function of the steady state reduced displacement potential(Ψ_{∞}), which is one source parameter. Table 1.1 relates typical physical parameters(rows) to

source parameters(columns).

A Comprehensive Test Ban Treaty(CTBT) emphasizes the identification of a nuclear explosion and thus the difference of these parameters relative those expected for chemical explosions and earthquakes becomes important. Relative comparison of parameters among several different explosions with almost identical initial conditions(similar geological and physical setting such as overburden pressure and source-material couple) might simplify the source modelling and make it easier to understand the physical background of the nuclear explosion, which is helpful in discriminating nuclear explosions from other sources.

High-quality seismic waveforms were obtained from free-surface instruments(less than 3 km from the source) and from the instruments installed at the same depth as the explosion (free-field) in the near-source region(several hundred meters) from three nuclear explosions - Misty Echo, Mineral Quarry and Hunter's Trophy - detonated at Rainier Mesa, Nevada Test Site(NTS). These data sets are dominated by source contribution in contrast to typical regional and teleseismic waveforms which are complicated by complex path effects. Free-field observations are believed to be surface wave free further simplifying the waveforms. The analysis of the data in this range is important in investigating near-field and surface wave effects since observations transition from the zone of near-field to the far-field and from body

Table 1.1 Parametric Relations by Mueller and Murphy

	Ψ_{∞}	$k(= 2\pi f_o)$
W	$W = c_1 \Psi_{\infty}^{1.32}$	$W = c_2 \frac{h^{1.26}}{k^3}$
r_{el}	$r_{el} = c_3 \frac{\Psi_{\infty}^{0.44}}{h^{0.42}}$	$r_{el} = c_4 \frac{1}{k}$
r_c	$r_c = c_5 \frac{\Psi_{\infty}^{0.38}}{h^{0.11}}$	$r_c = c_6 \frac{h^{0.26}}{k^{0.87}}$

Note: $c_1 \dots c_6$: medium dependent constants

W : yield

r_{el} : elastic radius

r_c : cavity radius

Ψ_{∞} : steady state reduced displacement potential

k : angular corner frequency

wave dominant to surface wave dominant. One of the goals of this study is to investigate biases of source parameters by the path effect. The free-field data, which is recorded from the seismometer buried under the ground, will be useful in investigating bias introduced by surface waves and the weathered layer when compared with the analysis of the free-surface data. The estimation of parameters for the purposes described above will be accomplished by non-linear inversion with bootstrap method.

Seismic waveforms can be expressed by the convolution of the source, path(Green's function), site effect and instrumental response in the time domain. For a mathematical treatment, each term can be expressed as a separate function or model.

$$(1-1) \quad u = S \otimes G \otimes W \otimes R$$

where S :Source function
 G :Green's function
 W :Site(weathered layer) effect
 R :Instrument response

Operator \otimes represents convolution. Each of these terms will be discussed in subsequent sections.

Source Models

There are two types of physical sources - explosions and earthquakes. For a nuclear explosion, there are various kinds of source models suggested both by theoretical considerations and empirical results (Haskell, 1967; von Seggern and Blanford, 1972; Mueller and Murphy, 1971; Helmberger and Hadley, 1981). Out of the existing explosion source models (Denny and Johnson, 1991), two types - Sharpe's and Haskell's - are generally used to represent the nuclear explosion sources.

The equation of motion in an elastic medium can be converted to the wave equation by introducing a displacement potential. Sharpe (1942) derived the displacement potential theoretically for an arbitrary form of pressure applied to the interior surface of a simplified spherical cavity. His functional representation of displacement potential contains a Fourier double integral.

$$(1-2) \quad \Phi = \frac{a}{2\pi\rho r} \iint \frac{p(\gamma) e^{in(\gamma-\tau)}}{n^2 + \frac{4in\alpha}{3a} - \frac{4\alpha^2}{3a^2}} dn d\gamma$$

where Φ : displacement potential
 a : cavity radius
 ρ : density of the medium
 r : distance
 p : pressure function

α : P-wave velocity

τ : retarded time = $t - \frac{r-a}{v}$

n, γ : Fourier index

Displacement can be derived by differentiating the displacement potential Φ with respect to the distance r . Mueller and Murphy (MM) followed Sharpe's derivation of displacement with an assumed pressure function and added some empirical relationships for different materials to derive their model (Mueller, 1969(a); Mueller and Murphy, 1971). Their model is expressed in the frequency domain as follows:

$$(1-3) \quad u(\omega) = \frac{P(\omega) r_{el}}{4\mu r} \frac{i\omega\alpha}{\omega_0^2 + i\omega_0\omega - \beta\omega^2}$$

where
$$P(\omega) = \int_0^\infty (\sigma_1 e^{-\omega_1 t} + \sigma_0) H(t) e^{-i\omega t} dt$$

r_{el} : Elastic radius

$u(\omega)$: displacement spectrum

r : Source-receiver distance

α, β : P- and S-wave velocity

ω_0, ω_1 : characteristic frequency

This model is proportional to ω^{-2} at high frequencies where $\omega \gg \omega_0$. Denny and Goodman (1990) derived ω^{-3} decaying source model from a slightly different pressure function.

Haskell's (HS), von Seggern-Blanford's (VSB), and Helmberger-Hadley's (HH) models are based on the same

theoretical background of displacement-reduced displacement potential(RDP) relationship for a spherically symmetric source. They used reduced displacement potential to remove the distance dependence on the displacement potential. The relation of the RDP and the displacement potential will be discussed in the subsequent section. Unlike Sharpe's type of source function which focuses on the physical pressure function at the source(Mueller-Murphy and Denny-Goodman), these other models(HS, VSB, HH) try to match seismic observation to simple polynomials. The difference between them is the order of the polynomials which are basically the form of Taylor expansion of the RDP function with a correction term and corresponding coefficients which are chosen to match the smoothness of the first motion in the observed data. Consequently, each model has a different order in its rise time which controls the high frequency roll-off(Figure 1.1 and 1.2) which is related to the continuity of the pulse(Savage, 1972). The VSB and HH model are based on the assumption that the intuitive Haskell's assumption of continuity of acceleration at the initiation time is not necessarily true in an explosion. And they neglected corresponding higher order terms. More details about these models will be discussed in Chapter 2.

Many complementary models exist for earthquake sources as well(Haskell, 1964; Aki, 1967; Brune, 1970). The main difference between the earthquake and the explosion in

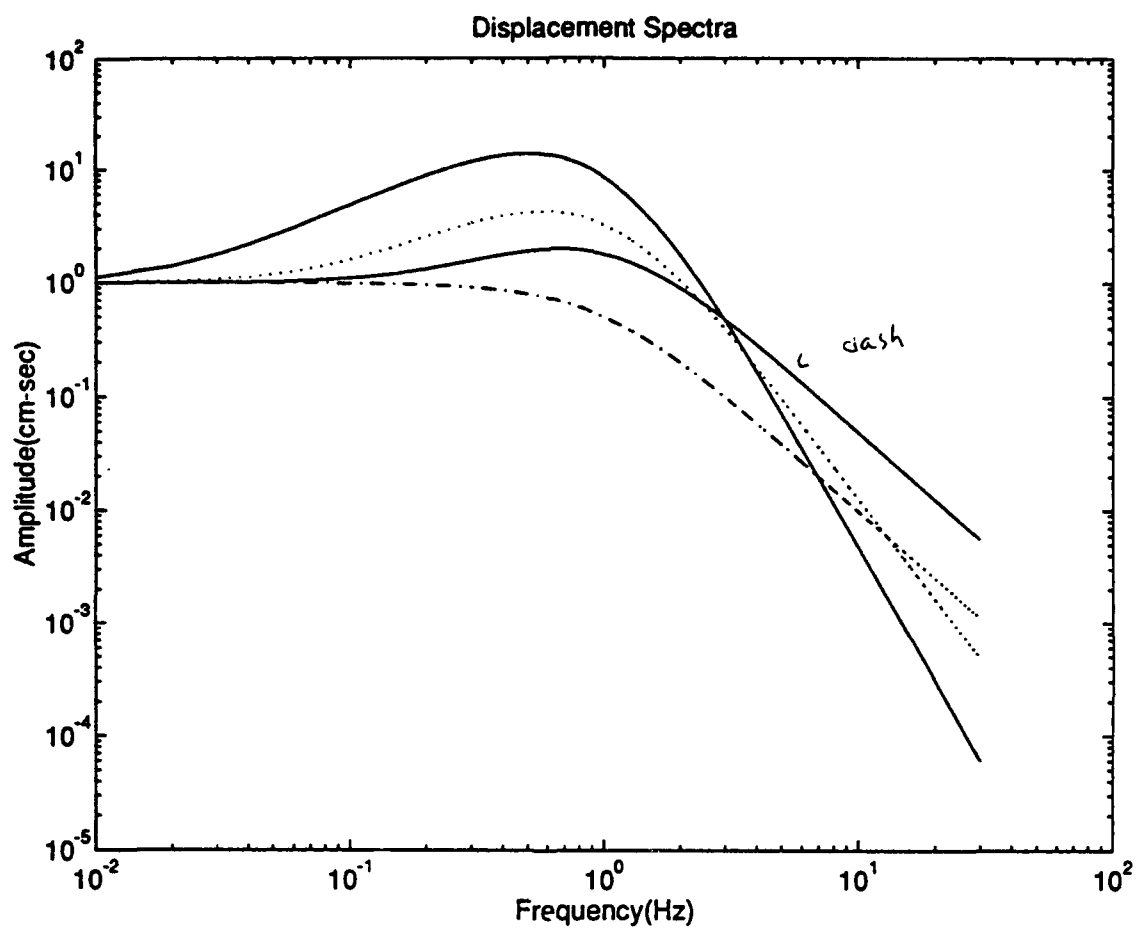


Figure 1.1 Displacement spectra from each source model. Long period spectral level is normalized. Solid, dotted, dashed and dash-dotted line denote HS(B=2, Fo=1), HH(B=2, Fo=1), VSB(B=2, Fo=1) and BR model(Fo=1) respectively.

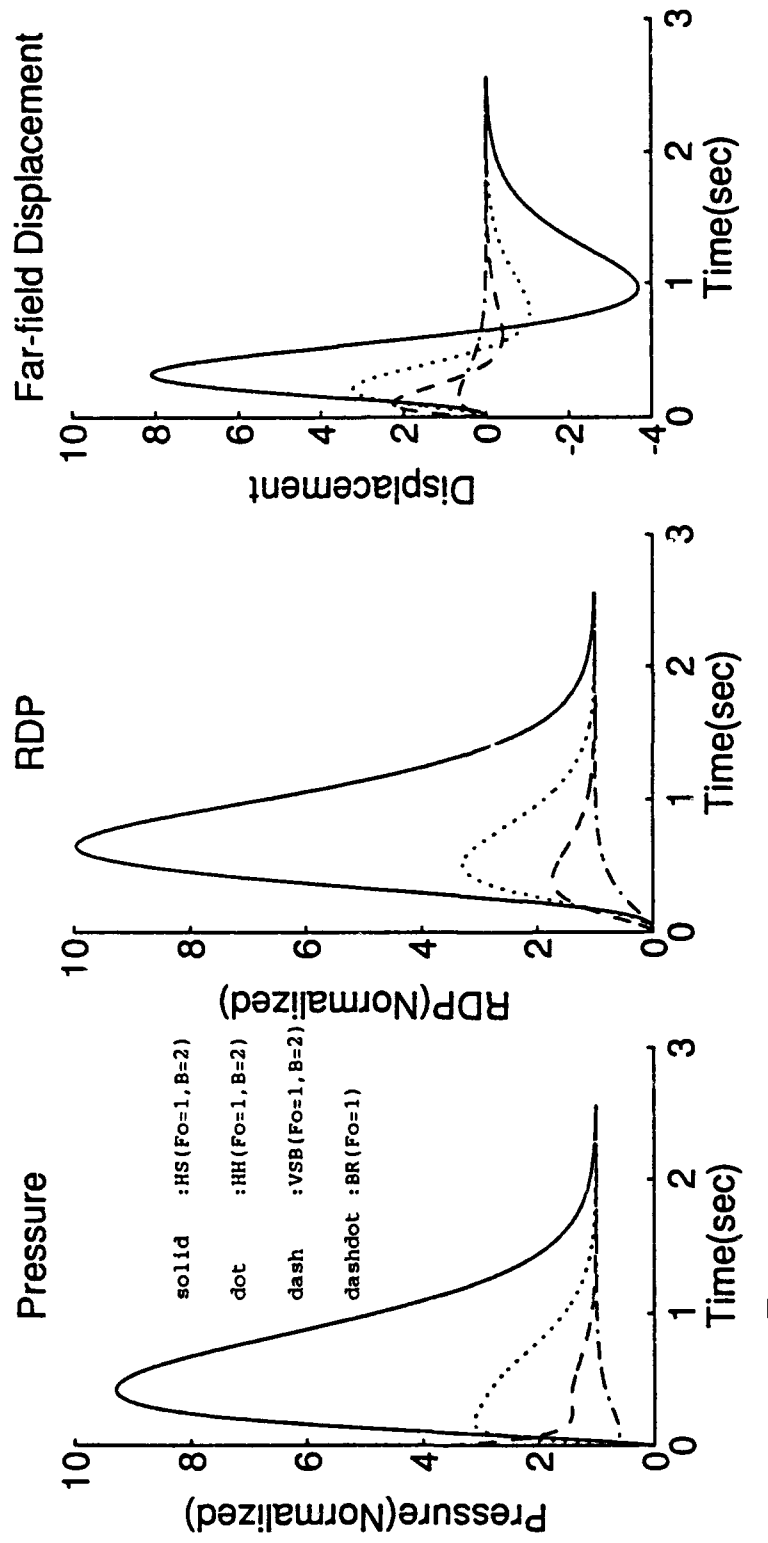


Figure 1.2 Pressure, RDP and displacement source time function from various models. Note that the definition of overshoot (B) changes as with the models.

modelling the source mechanisms is that the former is based on the planer motion while the latter is on the spherical volumetric motion and thus rich and poor in S-waves respectively. In our analysis, we will limit our consideration to P-wave source model since the explosion develops no S-waves theoretically even though there are some S-waves observed, possibly generated by the release of existing tectonic stress, by the conversion of waveform at layer boundaries and by spallation in actual observations. The other characteristic difference between the two is in their overshoots which are believed to be the result of rebound of the released energy. If the overshoot of the explosion is nearly zero, the earthquake source model can be used to model the shape of the explosion time function. One of the simplest earthquake models is Brune's model based on a stress drop across a fault as proposed by Brune(1970). This model has the same functional expression, as will be shown later, as the von Seggern-Blanford explosion model with no overshoot.

Even though each model is represented by different functional forms, they can all be expressed in terms of long period spectral level(LPL) or steady state RDP(Ψ_∞), corner frequency, overshoot and high frequency decay. Denny and Johnson(1991) summarized various existing explosion models extensively in the complex Laplace domain. These source parameters not only have physical meaning by themselves but

also are easy to formulate. Since most existing models are expressed by a few parameters, comparisons between the models can be derived (for example, von Seggern and Blanford, 1972; Denny and Johnson, 1991).

Long Period Displacement Spectral Level (LPL) and Steady State Reduced Displacement Potential (Ψ_∞)

The long period displacement spectral level (LPL) is defined as a flat level at long periods extending to the DC level (Figure 1.1) in the displacement amplitude spectrum (Brune, 1970). It is also called long period spectral level or long period flat level. It is proportional to a permanent displacement at a source. Since it is the residual radial displacement of the cavity produced by an applied pressure at the origin of the explosion, the LPL is proportional to some measure of source size, but it is a function of distance as well owing to geometrical spreading or wave propagation in complex materials.

Werth and Herbst (1963) introduced "reduced" displacement potential to analyze explosion data. It is defined by the relation with displacement.

$$(1-4) \quad u(r, t) = \frac{\partial \Phi}{\partial r} = \frac{\partial}{\partial r} \left[\frac{\Psi(t - \frac{r}{\alpha})}{r} \right]$$

where Φ is the displacement potential and Ψ is the reduced

displacement potential. Reduced Displacement Potential, unlike displacement potential, is independent of distance in a spherical coordinate system if homogeneous full-space structure is assumed. Though Reduced Displacement Potential is defined in the time domain and LPL is defined in the frequency domain, they are interrelated. When one thinks of only a far-field motion in a homogeneous full space, the scaling law of steady state Reduced Displacement Potential, LPL and moment in explosion (Müller, 1973; Aki et al, 1974; Denny and Johnson, 1991) is

$$\Psi_{\infty} \approx \alpha r \Omega_0$$

$$(1-5) \quad \Omega_0 = \frac{M_0}{4\pi\rho\alpha^3 r}$$

$$\Psi_{\infty} \approx \frac{M_0}{4\pi\rho\alpha^3}$$

where Ψ_{∞} : Reduced Displacement Potential at $t = \infty$

Ω_0 : LPL at frequency $=0$

M_0 : moment

α : P-wave velocity

r : distance

ρ : density

The relation between Reduced Displacement Potential function($\Psi(t)$) and steady-state Reduced Displacement

Potential(Ψ_{∞}), which is selected as a source representation parameter in this work because of its independence of time and distance, will be presented in Chapter 2.

Corner Frequency and High-frequency Roll-off

Corner frequency is the characteristic frequency beyond which spectral decay occurs. It is related to the duration of the displacement in the time domain(Figure 1.2) and is believed to be independent of the path. The corner frequency can be defined as the point at which two asymptotic lines of LPL and high frequency slope meet(Brune, 1970) in the displacement spectra. In earthquakes, a corner frequency can be uniquely related to the equivalent source radius(Brune, 1970) and the fault area(Savage, 1972), and thus the combination of the stress drop and moment(Anderson, 1983). In explosions, corner frequency is related to the elastic radius(Mueller and Murphy, 1977) and the yield of the explosion if we assume the spherical source model. It is proportional to the cube root of elastic source volume or an inverse of characteristic time(expansion time) for energy to travel from the center to the elastic boundary in which elastic waves start to propagate. Denny and Goodman(1990) have shown that the relation of corner frequency and the elastic source volume could be changed by assuming different pressure functions. Their calculations in the Laplace domain show that the corner frequency satisfying the Brune's

asymptotic definition does not fit the Mueller-Murphy's type of scaling law between the source pressure function (and thus yield) and the corner frequency. Regardless of its physical meaning or the relation with the source pressure function and yield, corner frequency can be treated as a mathematical parameter in an inversion since it shows characteristic distinction in the frequency domain.

The rate of high frequency decay beyond the corner frequency of HS, VSB, and HH models is dependent on the order of polynomial used to represent the rise in the source time function. The sharper the rise time, the greater the decay rate at high frequencies. The BR and MM models decay proportional to ω^{-2} beyond the corner frequency where ω is an angular frequency. For the HS, HH, and VSB models, the rate of decay varies according to the assumed order of RDP polynomial, ω^{-4} , ω^{-3} and ω^{-2} respectively. Physically, the HS model is the only one which is continuous in acceleration. There were some physical considerations related to the high frequency roll-off. Hanks and Wyss (1972) showed that the high frequency roll-off should be greater than 1.5 to be energy convergent in the far-field. Randall (1973) showed that the roll-off constant should be an integer, otherwise the velocity shows a branch-point singularity in the time domain. He also pointed out that ω^{-2} was the worst possible model since it showed finite discontinuity in velocity. If the absolute value of roll-off constant is less than 2, the

velocity shows infinite discontinuity at the origin time. Even though von Seggern and Blanford(1978) mentioned that there was an indication that the velocity was finitely discontinuous near the source, it is generally accepted that the velocity is continuous due to the elastic precursor which is generally observed at the very near-source range(within several hundred meters from the source).

Attenuation(Q) also affects the high frequency decay in the observed data, especially beyond the second corner frequency, f_{\max} defined by Hanks(1982). Burger et al(1987) estimated the yield and Q from far-field P-waves and found that trade-offs between the source model and anelastic attenuation exist.

Overshoot

One additional characteristic of the explosion source is the overshoot ratio which is the ratio of the peak RDP to the steady state potential. Mathematically, the coefficient of the correction term in each RDP polynomial of the source function determines the rate of overshoot. Physically, it represents the rate of rebound of the material to the applied pressure(von Seggern and Blanford, 1972) and is dependent of the medium around the source, but independent of the source depth or the yield(Denny and Johnson, 1991). Unlike others, Peppin(1976) introduced explosion source model without overshoot. It is also noteworthy that no apparent overshoot was reported in an early works in tuff(Werth and Herbst,

1963; Mueller and Murphy, 1971; Haskell, 1967; Peppin, 1976) which was generally interpreted as the non-elastic property(pore crushing) of the porous media. This is the material in which the explosions at Rainier Mesa are emplaced. There is a contradictory observation also. Denny and Goodman(1990) observed no apparent overshoots from the reevaluation of SALMON data. Based on this work, Denny and Johnson(1991) explained the possibility that porous media have a significant overshoot while non-porous media do not.

These three parameters - steady state RDP, corner frequency and overshoot - and a particular forward model can be used to characterize the explosion source; they can be obtained directly from the seismogram in the frequency domain and can be converted to various physical parameters such as static moment, yield, or elastic radius. The scaling relations of Mueller and Murphy(1971(a)) and Murphy(1977) are listed in Table 1.1.

Path Model and Attenuation

In order to complete source estimates, propagation path effects must be taken into account. They can be modeled simply by a geometrical spreading and attenuation effect for body waves. Homogeneous full-space, three-dimensional Green's function produced $1/r$ decay which is the simplest possible transfer function. The VSB and HH model is based on this assumption. Along with the homogeneous full-space, the homogeneous-half space(Johnson, 1974) and layered

structure(reflectivity method by Müller, 1985) can be considered in estimating the source parameters more accurately.

Energy attenuation of the traveling wave is due mainly to two physical processes - scattering and absorption. The scattering(scattering Q) occurs when the wave encounters an obstacle in an inhomogeneous media, while the absorption(intrinsic Q) is related to the non-elastic time-dependent property of the medium(Lomnitz, 1957; Stacey et al, 1975; Liu et al, 1976). Since Q is the total effect of these two physical processes, it is easier to define Q phenomenologically (Aki and Richards, 1980) rather than physically. It is well known that the wave should be dispersive to satisfy the causality, linearity and constant Q . For example, Azimi et al(1968) expressed acausal Q from the dispersive waves. Kjartansson(1979) derived linear acausal constant Q from the wave equation by assuming complex modulus. Though there are some arguments about the frequency dependency of Q (Minster, 1978 (a) and (b); Futterman, 1962), frequency independent Q is generally accepted in the frequency range of elastic waves in the earth(Knopoff, 1964; Liu et al, 1976). Berger et al.(1987) showed that there was no clear preference between the frequency-dependent Q model and frequency-independent Q model in matching observed narrow-band P-wave data. All of the above work is based either on theory or on far-field observations. In the near-

source region, McCartor and Wortman(1990) suggested the possibility of amplitude-dependent non-linear Q beyond the elastic radius by comparing calculations with the data obtained from 166 to 660 m from the nuclear source, SALMON. Non-linear Q can explain the precursors observed in the nuclear explosion seismogram without assuming complicated source physics. In this case, the concept of Q is not a robust description(Wortman and McCartor, 1991) since the source description is based on the concept of linear waves. This possibility, however, was excluded in this work since it is beyond the scope of this research. In this work, frequency independent Q was selected as an attenuation model for simplicity.

Spall

Spall, ballastic freefall of mass failed by the reflected tensile stress from the free surface, is a characteristic phenomenon in an explosion. The physical process, the generation of the elastic wave and the effect to the seismogram were extensively investigated by a number of researchers(Viecelli,1973; Stump, 1985; Taylor and Randall,1989). The source parameter estimation through an inversion can be biased by the spall effect mainly due to the frequency bandwidth of the spall(0.2 - 2 Hz, Tayler and Randall, 1989) which generally overlaps with the corner frequency of the explosion. This secondary source effect, however, was not considered here to simplify the problem.

Site Effect

It has long been known that the local site condition may be responsible for the fluctuation of spectral amplitude. By analyzing 75 accelerograms recorded in Italy from moderate to strong earthquakes ($4 \leq M_L \leq 7$), Rovelli et al (1988) found that there were significant differences in spectral shapes at different stations for the same earthquake. The degree of variation was greater than that obtained from the records of different earthquakes at the same site. Along with the fluctuation of the spectral amplitude and shape, it was recently reported that the site effect could cause directional resonance which amplifies the motion in one preferred direction leaving the others unaffected or diminished (Bonamassa and Vidale, 1991). Though it is not yet well understood, lateral inhomogeneity of very near receiver materials are the most likely candidate for these fluctuations and resonances. Coherence analysis by Menke et al. (1990) and Reinke and Stump (1991) show that there exists a variability in amplitude and phase even at closely spaced receivers installed at the same site. These variation are interpreted in terms of the inhomogeneous structure under the station. Unfortunately, the process of lateral inhomogeneity is stochastic rather than deterministic, which makes it hard to express as a functional form. In this work, the degree of fluctuation of source parameters through the entire recordings will be analyzed.

Instrument Response

Figure 1.3 shows nominal acceleration response of the accelerometers which were used to recover data from the nuclear explosions(MISTY ECHO, MINERAL QUARRY and HUNTER'S TROPHY). It shows flat and linear response up to about 100 Hz in amplitude and phase. Data Acquisition Systems(DAS's) by REFTEK can record the data up to 1000 samples per second with an antialias four pole Butterworth filter at 250 Hz. Since my intention is limited to several tens of cycles per second(Hz) due to the background noise, the system's response is sufficient for this purpose. Pearson(1992) presented detailed information and specifications for the instrument.

Purpose of the Work

Many different attempts have been made to calculate source parameters and the relationship among them. Burger et al.(1987) estimated the yield and Q from near-field and far-field seismogram and found that there was no preference for different source models if an appropriate attenuation model was selected. They compared the observed narrow band P-wave teleseismic data from the Pahute Mesa nuclear explosion with synthetic seismograms which were obtained from investigations of source and attenuation models using near-field data and concluded that there is no preference between ω^{-2} and ω^{-3} models in the frequency band of interest(0.5-4 Hz) if the frequency dependent Q model is used. It should be noted that

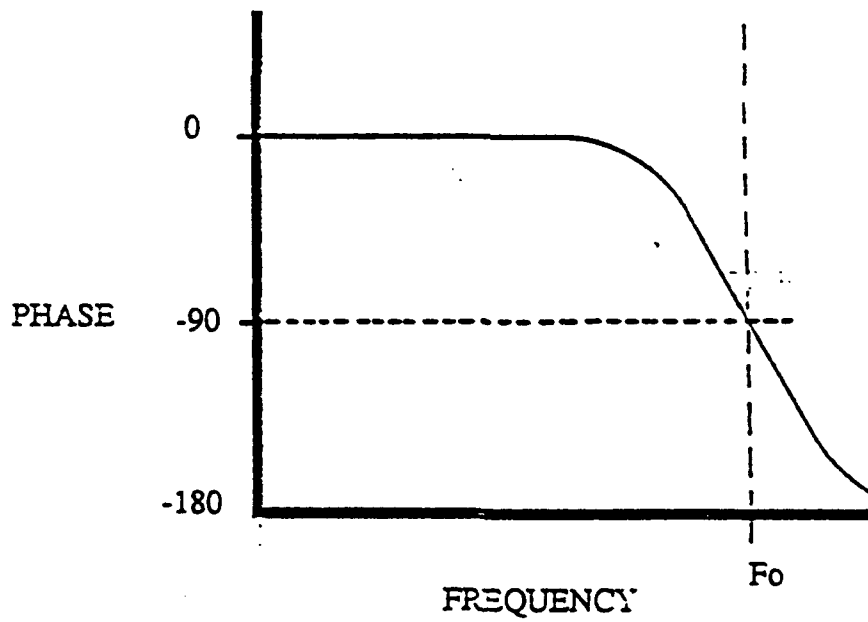
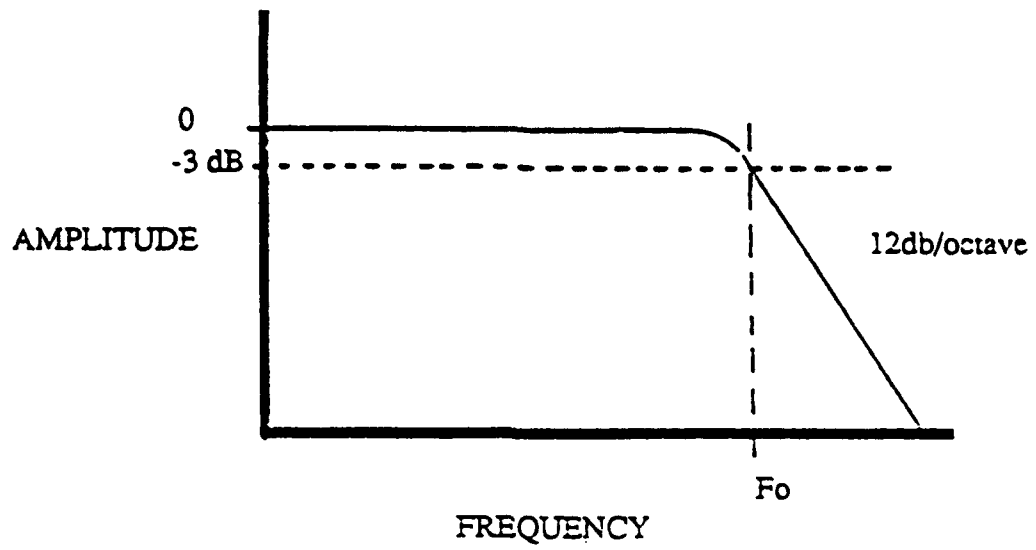


Figure 1.3 Accelerometer response curves.

the synthetic spectra generated from the t^* information by Der et al (1985) were treated as a near-field observation. The t^* is another way to express energy absorption along the ray path. It corresponds to the integral of Q^{-1} along the ray path. Hanks (1981) and Bakun and Bufe (1975) pointed out the relation between the corner frequency and anelastic attenuation. They emphasized that appropriate path correction should be made in order to estimate reliable source parameters. The same effect may be found in the analysis of explosion waveforms. The existence of S waves, anomalous surface waves, spall contributions and site effect which cannot be expected from the P wave model would bias the source parameters and Q in the explosion source model. One of the purpose of this work is to verify the degree of bias influenced by these secondary effects.

Inversion methods are powerful tools for characterizing observed data and resolving model parameters. By matching the data to the known forward model, one can get reliable estimates of model parameters assuming the model has been appropriately parameterized. This method is not only quantitative and objective but provides an opportunity to separate the effects such as source, path, site and instrument. Since Backus and Gilbert (1968) used this procedure in resolving the earth structure from the normal mode oscillations, inversion has been widely used in geophysics. Several attempts have been made to determine

source characteristics. Stump and Johnson(1977) used the linear inversion with singular value decomposition to represent the source as a set of moment tensors, which are the mathematical representation of the source, from a set of seismograms. This method is especially useful in separating non-spherical effects such as a spall in explosions. The main disadvantage of moment tensor inversion method is that it does not relate the source time function to the physical source parameters. Andrews(1986) used the inversion method to estimate the averaged event spectrum, and calculated the source parameters for Brune's earthquake model from event spectra. He calculated the mean value of amplitude at each frequency and estimated the source parameters. This procedure assumes that the path effect is randomized by an averaging process. His method would be useful where the path is not known. In both studies, a number of records were used to calculate the source parameters. On the other hand, since the recorded data always have many data points and the source can be explained by a few parameters, single records may be used in the inversion. Sereno et al(1988) and Scherbaum and Wyss(1990) used single station records to estimate parameters, but they used far-field data which were seriously affected by path effects. As a result, they emphasized the Q-structure rather than the source parameters.

For this work, BR(earthquake), VSB(explosion) and HH(explosion) models are used as forward source

representations. These models can be expressed relatively simply in the frequency domain and can be treated as representative of earthquakes, slowly decaying and rapidly decaying explosion models.

The purposes and goals of this work are summarized below;

(1) Define a method for recovering source parameters in a quantitative way. There are some difficulties to inverting the single spectrum for model parameters since

a) the inversion is non-linear

b) as frequency increases, the amplitude decreases rapidly and the inversion requires a weighting scheme

c) low frequency data before the corner frequency are easily contaminated by the noise and the noise biases the estimation of source parameters seriously

(2) Explore the trade-offs between model parameters. It is well known that each source parameter such as Ψ_∞ , overshoot and corner frequency exhibit trade-offs. It will be important to verify the reason for these trade-offs and the effect to the parameter estimates if one or more parameters are constrained.

Burger et al's (1987) analysis of the trade-off between the high frequency decay-rate and Q will be tested using very near-source data, including free-field data.

(3) Quantify the effect of propagation path at small ranges.

Near-source studies can be useful in characterizing the source because the recorded data are less contaminated by complicated propagation path effects. Three kinds of propagation path corrections will be tested. They span from the simplest and easiest to apply to the most complex.

- a) homogeneous full-space path
- b) homogeneous half-space path
- c) layered structure

It will be checked whether a simple path model is sufficient to extract reliable source parameters from the recorded data. The applicability and comparison of these path corrections will be analyzed with a combination of near-field, far-field and surface wave effects using synthetic and recorded data. Scattering by the lateral heterogeneity, spall and local site effects will not be considered. Anisotropic, inelastic and non-linear wave propagation will not be considered either.

(4) Compare the applicability of source models. Three different source models will be used as a forward model of the inversions;

- a) BR(Brune's) model
- b) VSB(von Seggern-Blanford's) model
- c) HH(Helmberger-Hadley's) model

Source parameters obtained from each model with the synthetic data and near-source data will be compared.

(5) Compare the source estimates from free-field and free-surface data. This will help quantify the local site effect,

especially by weathering which is expected to be quite variable from one free surface receiver site to another.

(6) Verify the factors which bias source estimates, including the effects of surface waves, near-field terms, site response and the model parameterization.

(7) Estimate the statistics of source parameters from different explosions at the same site and compare them. In order to obtain the reliability from the scattered estimates by the effects which are not considered in formulating the model, bootstrap estimation will be performed.

CHAPTER 2

INVERSION

The forward model will be introduced and discussed in this chapter. The different source models will be transformed into the model space. The similarities and differences among them will be presented along with the path models used for the correction of the data.

Invertibility and regularization are important in geophysical inversion. The source of ill-posedness and large condition number, and the trade-off between the parameters will be discussed. A regularization method will be presented. The development of several inversion schemes will be introduced. They will be compared to each other using the synthetic data and an optimal model will be selected. A modified scheme which combines several techniques maintaining an advantage of an optimal model will be tested.

The procedure for simultaneous inversion for source parameter estimation will be introduced. Simultaneous inversion with data from different stations may strengthen the assumption of randomness of the data in the application of least square method.

Norm based inversion has limitation in its applicability in the source parameter inversion as will be illustrated in the early part of this chapter. The steady state RDP shows ill-posedness since the data before the corner frequency is quite limited. It is also easy for the low frequency amplitudes to be contaminated by noise due to the shape of the source acceleration amplitude spectra. Broad-band data will solve the ill-posedness of the steady state RDP in the source parameter inversion in the frequency domain, but it is not available in the near-source region. The other possibility to increase the robustness of the estimates (especially for the steady state RDP and corner frequency estimation) is use of the bootstrap method. The basic theory of non-parametric bootstrapping and its application in the source parameter inversion will be presented.

Forward Model

As was mentioned in Chapter 1, the source time function for an explosion can be separated into two groups; Sharpe-Blake type and Haskell type. The Haskell type of source model was used in this work since it is easy to formulate as a parameter expression. The von Seggern-Blanford and Helmberger-Hadley models were adopted for a test. The forward model can be separated into three parts - source function, Green's function and attenuation function - in a

source parameter inversion if we do not consider the secondary-source spall, fluctuation by lateral heterogeneity and site effect. Three different types of path models - homogeneous full-space, homogeneous half-space, and layered structure model - were used as a path model. Frequency independent Q model was used as an attenuation model.

Source

Haskell's model

Based on the shape of the experimentally determined RDP's of Werth and Hurbst(1963), Haskell(1967) introduced a simple analytic RDP parameterization.

$$(2-1) \quad \Psi(t) = \Psi_{\infty} \left\{ 1 - e^{-kt} \left[1 + kt + \frac{1}{2}(kt)^2 + \frac{1}{6}(kt)^3 - B(kt)^4 \right] \right\}$$

where Ψ_{∞} : reduced displacement potential at $t \rightarrow \infty$
 B : dimensionless correction term to be determined by the data
 k : constant to be chosen to fit the observed data
 t : retarded time

B , dimensionless correction term to be determined by the data, is interpreted as an overshoot and k , constant to be chosen to fit the observed data, is interpreted as a corner frequency. Since this RDP function is analytic with respect to time, it is continuous in displacement, velocity and

acceleration. In spite of its physical suitability, this model was excluded from the study due to the fact that it generally does not fit the slope of the rise time of the observed waveform which is linked to the high-frequency decay (von Seggern and Blanford, 1972). Although Haskell's model is not formally used in this analysis, its form was used as a basis in the functional development of other models such as the VSB and HH models that are used in this study. As a reference, the far-field displacement spectrum in the frequency domain can be obtained by substituting Equation 2-1 into the RDP-displacement relation (Equation 1-4) and transforming into the frequency domain.

$$(2-2) \quad |u_{ff}(f)| = \psi_{\infty} \frac{1}{\alpha \cdot r} \frac{\left[1 + (1 + 24B)^2 \frac{f^2}{F_0^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_0^2} \right]^{5/2}}$$

where $u(t)$: displacement
 α : P-wave velocity
 r : source-receiver distance
 F_0 : corner frequency

Comparison of the von Seggern and Blanford's model with Brune's model

For a spherically symmetric source, the relation between reduced displacement potential and displacement (von Seggern and Blanford, 1972) in a homogeneous full space is

$$(2-3) \quad u(t) = \Psi_{\infty} \left\{ \frac{1}{\alpha \cdot r} \frac{d [\Psi(t) / \Psi_{\infty}]}{dt} + \frac{1}{r^2} \frac{\Psi(t)}{\Psi_{\infty}} \right\}$$

where $u(t)$: displacement
 α : P-wave velocity
 r : distance from the source to the receiver
 Ψ_{∞} : reduced displacement potential at $t \rightarrow \infty$
 $\Psi(t)$: reduced displacement potential
 t : retarded time

The first term in the bracket which is proportional to $1/r$ is the far-field term while the second term which is proportional to $1/r^2$ is the near-field term.

von Seggern and Blanford assumed a functional form for the reduced displacement potential as

$$(2-4) \quad \Psi(t) = \Psi_{\infty} \{ 1 - e^{-kt} [1 + kt - B(kt)^2] \}$$

where k and B are values designed to fit the model to observed teleseismic short period P-waves from Amchikta Island explosions. Comparison to Equation 2-1 illustrates that they dropped higher order terms which yield slower rise time. It should be noted that the steady state RDP(Ψ_{∞}) are time-independent while the reduced displacement

potential($\Psi(t)$) is time-variant. Steady state RDP(Ψ_∞) will be used as a source representation parameter in this work and presented as Ψ_∞ or RDP. Time-variant reduced displacement potential will be presented as $\Psi(t)$ or RDP function to prevent confusion.

If Equation 2-4 is substituted into Equation 2-3, the displacement is expressed by

$$(2-5) \quad u(t) = \Psi_\infty \left(\frac{1}{\alpha r} (-k^2 t e^{-kt}) (2B+1-Bkt) \right. \\ \left. + \frac{1}{r^2} \{1 + e^{-kt} [1 + kt - B(kt)^2]\} \right)$$

If the Fourier transform is taken of Equation 2-5, after a straightforward but tedious calculation, the displacement spectrum becomes

$$(2-6) \quad |u(f)| = \Psi_\infty \left(\frac{1}{\alpha r} + \frac{1}{2\pi f r^2} \right)^{1/2} \frac{\left[1 + (1+2B)^2 \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{3/2}}$$

where $F_o = k/(2\pi)$ (corner frequency).

Velocity can be expressed by the differentiation of Equation 2-5 with respect to time in the time domain or multiplication of $(-i2\pi f)$ in the frequency domain (Equation 2-8).

$$(2-7) \quad |v(f)| = \Psi_{\infty} \left(\frac{4\pi^2 f^2}{\alpha^2 r^2} + \frac{1}{r^4} \right)^{1/2} \frac{\left[1 + (1+2B) \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{3/2}}$$

As distance increases, the higher order term($1/r^2$ in Equation 2-6 or $1/r^4$ in Equation 2-7) becomes insignificant, which leads to the far-field displacement(or velocity) spectra

$$(2-8) \quad |u_{ff}(f)| = \Psi_{\infty} \frac{1}{\alpha r} \left\{ \frac{\left[1 + (1+2B) \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{3/2}} \right\} \quad \text{and}$$

$$|v_{ff}(f)| = \Psi_{\infty} \frac{2\pi f}{\alpha r} \left\{ \frac{\left[1 + (1+2B) \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{3/2}} \right\}$$

The representation of displacement spectrum in Equation 2-8 is the same as Equation (12) in von Seggern and Blanford's(1972).

Equations 2-5, 2-6 and 2-7 can be interpreted as the multiplication of the source strength term(Ψ_{∞}), source shape term(curly bracketed terms on the right hand side of Equation 2-8 and equivalent terms in 2-6 and 2-7) and the full-space path term(distance dependent terms). Therefore, the full-space path term can be replaced by a more realistic path term such as homogeneous half-space or layered structure term.

There is no difference between the displacement and the velocity source terms.

If there is no overshoot ($B=0$), the VSB model becomes;

$$(2-9) \quad |u(f)| = \Psi_{\infty} \left(\frac{1}{\alpha \cdot r} + \frac{1}{2\pi f r^2} \right) \frac{1}{1 + (f/F_0)^2}$$

The Brune's model (1970) is expressed as;

$$u(f) = \frac{\Omega_0}{(f/F_0)^2 + 1}$$

where $u(f)$: displacement amplitude spectrum

Ω_0 : long period spectral level

F_0 : corner frequency

Comparing above equation to the VSB model (Equation 2-9) in a homogeneous full-space, we found the relation between the two models as follows;

$$\Omega_0 = \Psi_{\infty} \left(\frac{1}{\alpha r} + \frac{1}{2\pi f r^2} \right) \quad (r \neq 0)$$

$$B = 0$$

This relation shows that both models can be represented as the same source time function even though the source mechanism may be different. It is to be noted that Ψ_{∞} is distance independent while Ω_0 is dependent on distance.

If $-1 \leq B < 0$, the B value does not represent overshoot but it represents the slope of high-frequency decay rate(Figure 2.1). It should be noted, however, that the only value to be taken as B between this interval is $B=-0.5$ to maintain physical plausibility of integer roll-off in the high frequency range(Hanks and Wyss, 1972; Randall, 1973). If $B=-0.5$, the von Seggern-Blanford model reduces to the Helmberger-Hadley model without overshoot(shown in a subsequent section). It is reasonable to confine B as positive value.

There is a difference between the corner frequency defined in earthquakes and in explosions. The definition of k (or $f_{VSB}=k/2\pi$) in the VSB source model, like the other two similar models, is based on the time domain source function. It was chosen to fit the model to the rise time of the observed first wave motion. This definition is easily related to the corner frequency of Brune's in the frequency domain if there is no overshoot(defined by the two asymptotes as illustrated in Figure 2.2). If overshoot exists, however, f_{VSB} defined by the VSB model behaves differently from that of Brune's corner frequency. The VSB source spectrum with overshoot of 0.366 produces a corner frequency estimate, 0.62, which differs from the definition of k or $2\pi f_{VSB}$, 0.5(Figure 2.2). The approximate relation between the two can be expressed by $f_{VSB} \approx \frac{f_{BR}}{\sqrt{1+2B}}$ where f_{BR} is the corner frequency of the Brune model and B represents the overshoot.

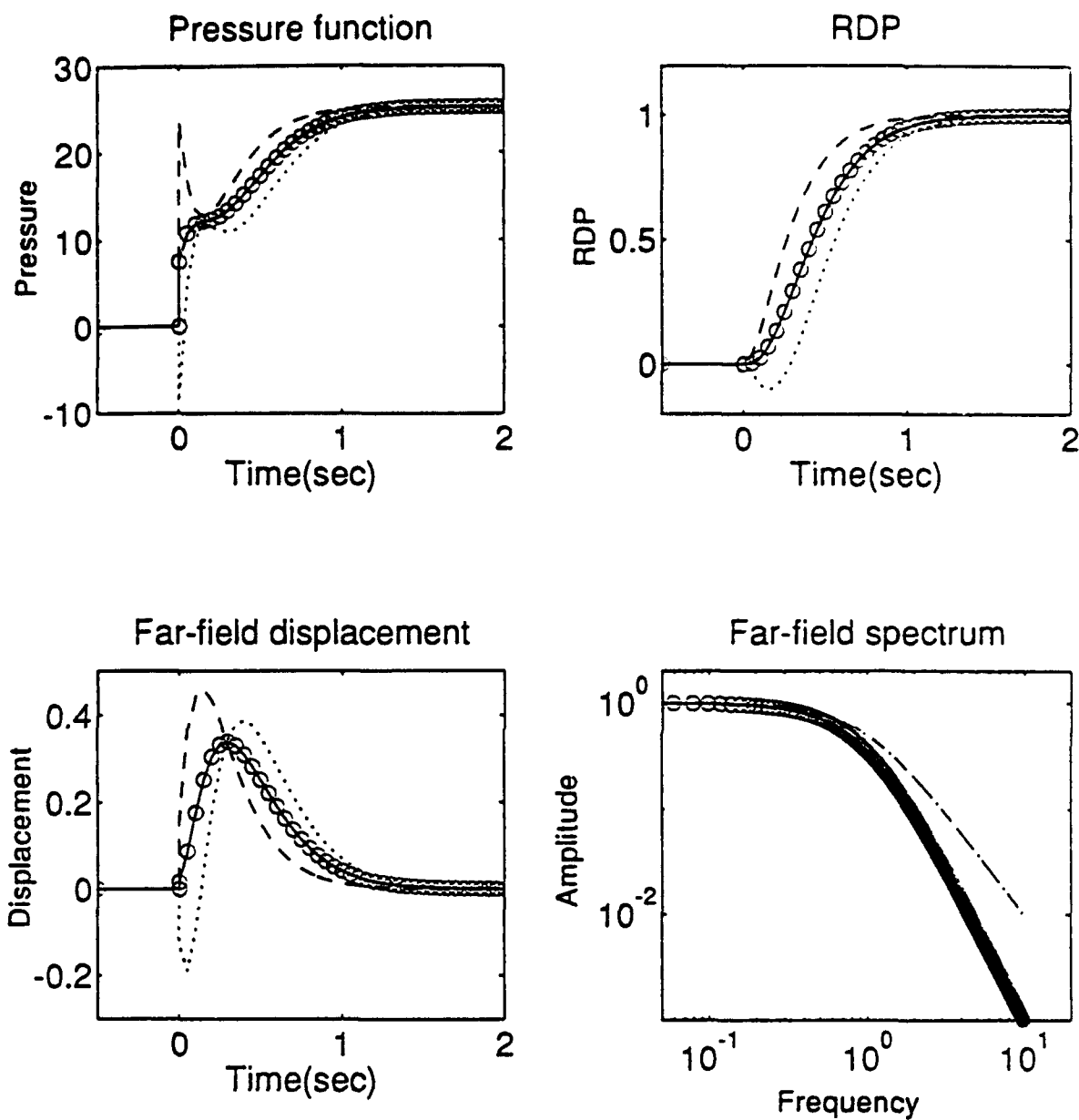


Figure 2.1 Source functions with various values of B in the VSB model. Solid line, dotted line and dashed line denote $B=-0.5$, $B=-1$, and $B=0$ respectively. Circle denotes HH model without overshoot.

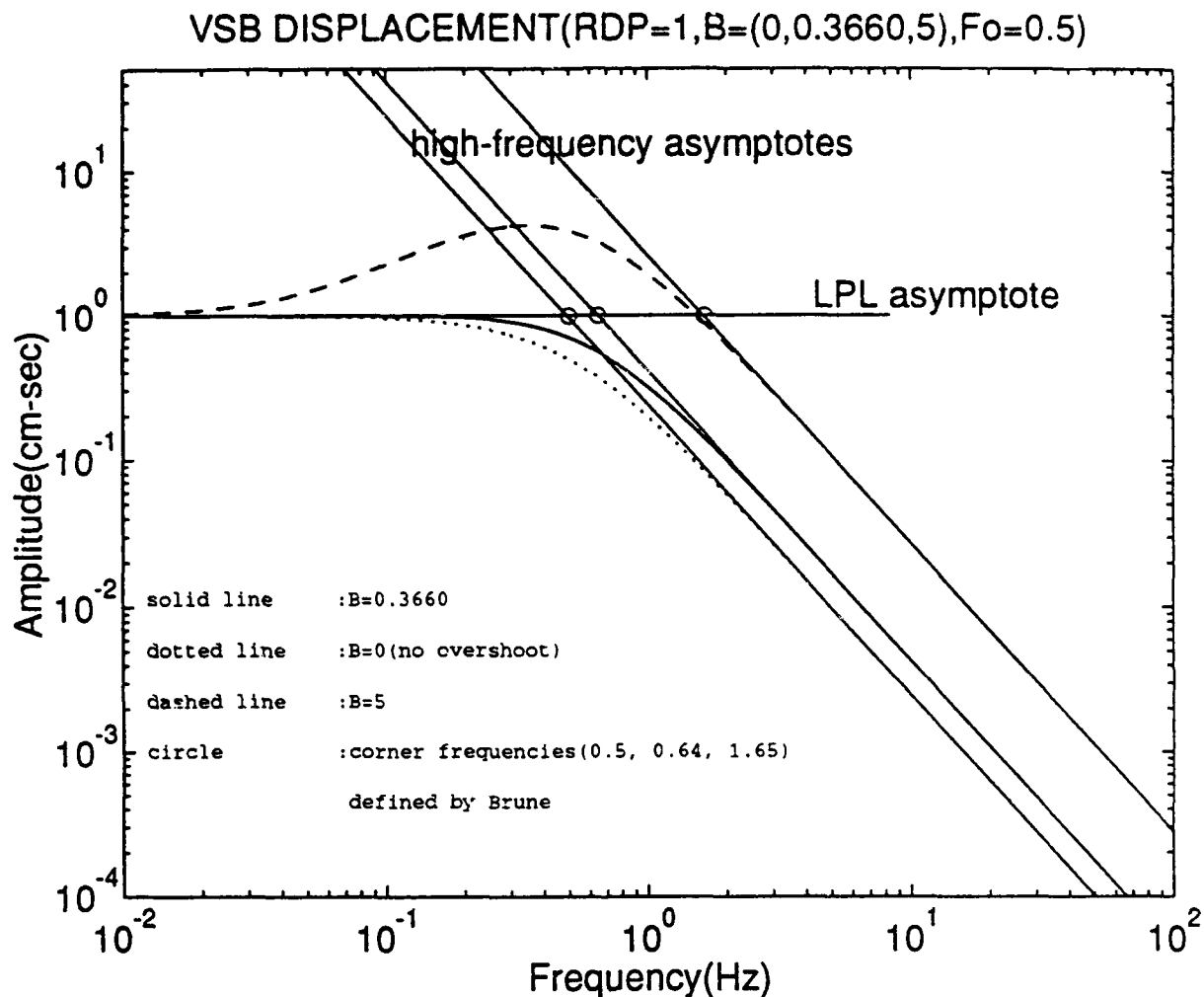


Figure 1.3 The relation of corner frequency defined by Brune and by VSB. When there is no overshoot, both definition agrees. If overshoot exists, corner frequencies (0.64 and 1.65) defined by Brune with two asymptotes are different from that defined by VSB (0.5). When the overshoot is high ($B=5$), it is even hard to define LPL asymptote.

It is to be noted that the corner frequency defined by Brune is significantly biased by overshoot. The detailed derivation of this relation is given in Appendix A. In Figure 2.2, RDP is the steady-state reduced displacement potential which is the same as LPL after normalization by distance and velocity, B is the overshoot ($B \geq 0$) and $F_{VSB}(=k/2\pi)$ is the corner frequency as defined by von Seggern and Blanford. Related to the above corner frequencies defined by the source models, the term f_{app} (apparent corner frequency) will be used in this paper. Apparent corner frequency is nothing but a biased corner frequency when the data with overshoot is interpreted by the Brune's model. It is natural that f_{app} lies within f_{VSB} and f_{BR} . Apparent corner frequency, therefore, is dependent on overshoot as well as on the method used to resolve the corner frequency. Apparent corner frequency not only reduces the confusion between the definition of corner frequencies but also can be recovered to the corresponding corner frequency if B value is known. Detailed derivation is in Appendix A.

Helmberger-Hadley's model

To match the limited number of near-field data (8 km from the source) from Pahute Mesa explosion, Helmberger and Hadley (1981) modified Haskell's representation.

$$(2-10) \quad \Psi(t) = \Psi_{\infty} \left\{ 1 - e^{-kt} \left[1 + kt + (kt)^2/2 - B (kt)^3 \right] \right\}$$

This equation is interemediate between the VSB and HS models. Therefore, it is continuous in displacement and velocity, but piecewisely discontinuous in acceleration. If we apply the RDP-displacement relation(Equation 1-3) and transform it into the frequency domain, the displacements become:

$$(2-11) \quad u(\omega) = \left(\frac{i}{\alpha \cdot r} + \frac{\omega}{r^2} \right) \left[\frac{(1 + 6B) k^3}{(\omega - i k)^3} + \frac{6B i k^4}{(\omega - i k)^4} \right]$$

where i and ω denote $\sqrt{-1}$ and angular frequency respectively. Neglecting the near-field term and converting angular frequency(radian/sec) to frequency(Hz), the amplitude spectra in the far-field become;

$$(2-12) \quad |u_{ff}(f)| = \psi_{\infty} \frac{1}{\alpha \cdot r} \frac{\left[1 + (1+6B)^2 \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{4/2}}$$

$$|v_{ff}(f)| = \psi_{\infty} \frac{2\pi f}{\alpha \cdot r} \frac{\left[1 + (1+6B)^2 \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{4/2}}$$

These results illustrate that B has a different representation from model to model.

Generalization

The Haskell type functions(VSB, HH, and HS) can generally be expressed by

$$(2-13) \quad \Psi(t) = \Psi_{\infty} \left\{ 1 - e^{-kt} \left[1 + kt + \frac{1}{2}(kt)^2 + \dots - B(kt)^n \right] \right\}$$

in the time domain and

$$(2-14) \quad |u_{ff}(f)| = \Psi_{\infty} \frac{1}{\alpha \cdot r} \frac{\left[1 + A^2 \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{(n+1)/2}}$$

in the frequency domain

where $A = 1 + c \cdot B$

$B =$ overshoot

$c = n!$ and

$n =$ the highest power in $\Psi(t)$.

The generalized form of the Haskell type source function is useful in the formulation of the inversion programming. The above representation makes it possible either to set decay-rate(n) as a parameter or to set it as constant for various source models.

Path

Three different types of path models are assumed in this work. These path models are used as correction terms, thus no parameters to be inverted are involved.

Homogeneous full-space model

This is the simplest path model used in this work. It is included in the VSB or HH source model. In this simple

case, more complex path effect would be mapped into the source parameters. The homogeneous full-space path model will be used for two purposes. The applicability of full-space path model to the source parameter inversion will be tested. Despite the apparent bias of source parameters by an inappropriate path correction, this path correction may be acceptable if the degree of bias is not serious in the near-source region. Synthetic seismograms from an elastic half-space will be inverted with this path model in the following section as an initial test of possible bias.

Free-field data were obtained from very near-source instrumentation in tunnels at the same level as the source. The path from the source to the buried receiver can be treated as a full-space homogeneous path since the direct wave effect is assumed to be dominant in this range.

Homogeneous half-space model

Simple homogeneous full-space path model can be replaced by the homogeneous, elastic half-space path model. Johnson(1973) derived analytic Green's function in an elastic homogeneous half-space from the equation of motion(Lamb's problem). The equations of motion and boundary conditions are transformed into the Laplace domain, solved and then transformed back into the time-space domain by the Cagniard-de Hoop method. The advantage of transformation is that the complicated differential equations can be manipulated

algebraically in the transformed domain. Since this solution is analytic and exact, it has all wavefield contributions including wave conversions, diffractions, surface waves, free-field interactions and near-field effect as well as the direct and reflected waves. Unfortunately, the solution is expressed as an integral equation which cannot be solved analytically. Numerical calculation of this integral equation may introduce error. This error tends to increase as the take-off angle approaches $\pi/2$ (shallow source depth relative to observation range).

Layered structure(Reflectivity method)

There are several ways to generate Green's function in a layered or realistic structure(Helmberger, 1968 for the Generalized Ray Theory; Cormier and Richards, 1977 for the Full Wave Theory; Chapman, 1978 and 1985 for the WKBJ Method; Fuchs and Muller, 1971, Muller, 1985 for the Reflectivity Method; Panza, 1985 for the Modal Summation Method). For the purpose of this study, examining the biases of source parameters due to various non-source effects, the reflectivity method was used for laterally homogeneous multiple layers. This method is not only appropriate for the whole wavefield(body waves, surface waves, multiples and near-field effects, etc) but is also numerically stable at all frequency and slowness ranges. The reflectivity method is also called the frequency domain method(Mooney, 1983) or

wavenumber(or slowness) integration method(Müller, 1985) since the Green's function is calculated in the frequency domain by summing up the reflectivity and transmissivity coefficient at each boundary(boundary condition) in the whole slowness range recursively and is transformed into the time domain. Ungerer's reflectivity program(1990) was used to generate the Green's function for the layered structure in this work. His reflectivity program is based on Müller's paper(1985) and includes near-field term(mid-range term in the expression of moment tensor) and calculates Bessel function analytically for the slowness summation.

Attenuation

If the exponentially decaying frequency independent Q model(Aki and Richard, 1980) is applied to Equation 2-7 and 2-12, since the convolution in the time domain is equivalent to the multiplication in the frequency domain, the displacement and velocity in a homogeneous full space become

$$(2-15) \quad |u_{ff}(f)| = \psi_{\infty} \frac{1}{\alpha r} \frac{\left[1 + (1+2B) \frac{f^2}{F_o^2}\right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2}\right]^{3/2}} e^{-\pi f t / Q} \quad \text{and}$$

$$(2-16) \quad |v_{ff}(f)| = \psi_{\infty} \frac{2\pi f}{\alpha r} \frac{\left[1 + (1+2B) \frac{f^2}{F_o^2}\right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2}\right]^{3/2}} e^{-\pi f t / Q}$$

for BR(B=0) and VSB model and

$$(2-17) \quad |u_{ff}(f)| = \Psi_{\infty} \frac{1}{\alpha \cdot r} \frac{\left[1 + (1+6B)^2 \frac{f^2}{F_o^2}\right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2}\right]^{4/2}} e^{-\pi f t / Q}$$

$$(2-18) \quad |v_{ff}(f)| = \Psi_{\infty} \frac{2\pi f}{\alpha \cdot r} \frac{\left[1 + (1+6B)^2 \frac{f^2}{F_o^2}\right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2}\right]^{4/2}} e^{-\pi f t / Q}$$

for the HH model. The differences between the VSB and the HH model are the expression of overshoot (1+2B for VSB and 1+6B for HH) and the power of the denominator (3/2 for VSB and 2 for HH) which controls the high frequency decay rate.

Inverse Process

Inversion is intended to extract information about the theoretical model from the observed data. The information to be determined is generally expressed as a parameter or a set of parameters. The parameter can be a number or a function. Parameter estimation, or inversion, has some concerns that must be investigated (Parker, 1977; Koch, 1992).

Existence of the solution is the fundamental property in the formulation of inversions. If the forward model or operator can be expressed continuously in Hilbert space (Koch, 1992), which is a complete inner product space or an infinite dimensional vector space where unit vectors (eigenfunctions or

basis) are orthogonal (Arfken, 1970; Backus and Gilbert, 1976), the model has a solution. This generalization of the existence of the solution is related to the Fredholm alternative. Assume that the linear model is expressed as

$$(2-19) \quad d = g(m, x)$$

such that the data are related linearly to the model parameters and variables through the operator g . In this case, g is called a linear operator mapping the model parameters from the model space to the data in the data space. It should be noted that the forward model is called linear (or non-linear) with respect to the model parameters (m), not with respect to the variables (x). In the non-homogeneous equation (Equation 2-19), the solution does not exist when there exists non-trivial homogeneous solution or null vector (Fredholm alternative) and when the non-trivial solution is not orthogonal to the operator.

$$\int d \cdot \phi \, dm \neq 0$$

where ϕ is the solution of the homogeneous equation $d(m, x) = 0$. It is, however, of little significance for the geophysical inversion since the solution will be found by the numerical method in most cases (Koch, 1992). Singular value decomposition is a powerful tool to find the best solution numerically by setting redundant small eigenvalues to zero.

In the above non-homogeneous equation(Equation 2-19), there may exist an infinite number of solutions if there exists a non-trivial homogeneous solution and if the operator and the non-homogeneous solution are orthogonal.

$$\int d\varphi \, dm = 0$$

This situation corresponds to one in which one of the eigenvalues in g is zero(Habermann, 1989). Non-uniqueness in a geophysical inversion was emphasized by earlier authors(Backus and Gilbert, 1967 and 1968; Parker, 1977) because of the finiteness of the data space and the infinite dimensionality of the model space. It is inevitable when one tries to represent the continuous model with a discrete data. Non-uniqueness is crucial when the forward model cannot be expressed as a functional. Fortunately in physics and mathematics, there are many cases when the forward model can be expressed as a mathematical function with a finite number of parameters with sufficient accuracy while data can be collected nearly continuously. In this case, the deviation of the data from the constructed forward model will be treated as theoretical noise(Tarantola, 1987). When the forward model can be expressed by a discrete finite function, identifiability will be more important than the uniqueness in the theoretical viewpoint(Koch, 1992). Identifiability is a possible inherent property of non-uniqueness in the discrete

forward model. The density structure in the gravitational anomaly inversion and Wiechert-Herglotz inversion in the seismology are typical example of identifiability. It is well known that the size of the anomalous body and density cannot be determined at the same time in the gravity anomaly analysis. The velocity cannot be identified from the regular travel time curve if the low velocity layer is included and the Wiechert-Herglotz inversion is thus non-unique.

Nowadays, the non-identifiable property between the velocity and thickness in a structure or between the source location and the velocity structure is the main issue in most tomographical inversions. If the forward model is not identifiable, inversion shows trade-off between parameters. Additional information, thus, is necessary to solve the non-identifiable problem in an inversion.

Along with the non-uniqueness and existence concern, there is an extremum problem in the non-linear inversion. The way to find the solution in the non-linear iterative inversion is to find the minimum error. There is no known way to figure out whether the calculated minimum is local or global.

The other practical and the most important concern in the geophysical inversion is stability. The system is unstable or ill-posed if a small perturbation of the data results in a large variation in the estimates. This is especially important when dealing with the numerical problem

since instability induces non-uniqueness to the model (Koch, 1992). The stability will be discussed in a subsequent section.

Characteristic of Forward Model

Existence and uniqueness

As was mentioned in the previous section, existence and non-uniqueness in an underdetermined case is not so significant in an application of an inversion in physics. Most significant concern related to these basic properties of mathematics might be the non-uniqueness by the identifiability or trade-off. The resolution matrix, or resolving kernel (Backus and Gilbert, 1968; Herrmann, 1985), identifies how well the model parameters are resolved. Uniqueness can be verified by comparing the resolution matrix to the identity matrix after inversion (Jackson, 1972). Refer Appendix B for more explanation.

Stability and regularization

The inverse process tends to be singular or nearly singular in some cases. Condition number, indicative of the singularity, is defined by the ratio of maximum and minimum eigenvalues. Geometrically, large condition number is characterized as nearly parallel sets of eigenvectors (Horn and Johnson, 1985). Figure 2.3 shows it graphically (Gerald, 1978). If the data is exact, the solution is unique and exact (a and c) regardless of its colinearity, unless it is perfectly colinear. If two lines (eigenvectors) are

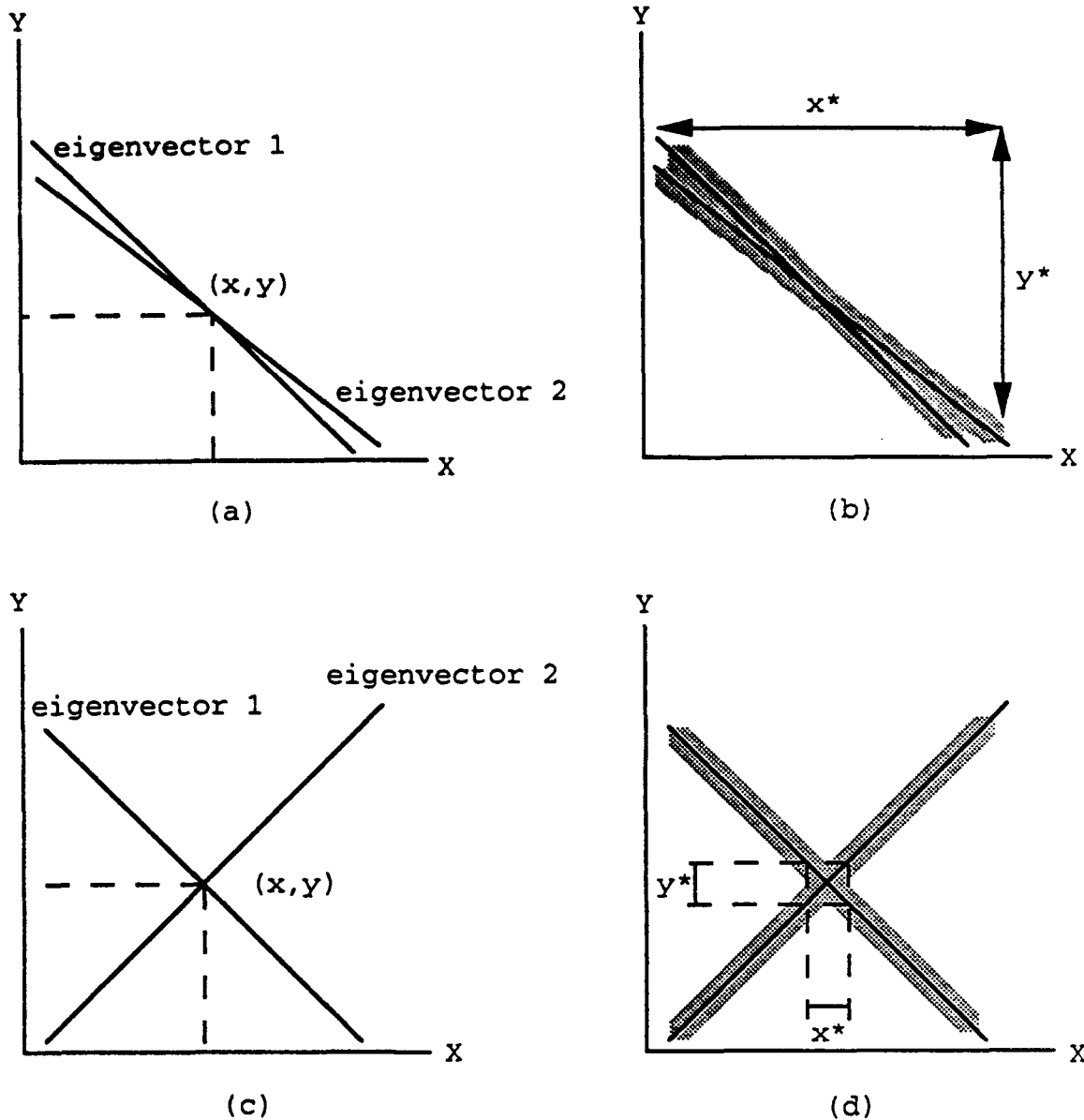


Figure 2.2 Graphical representation of colinearity of eigenvectors. If there is no error, unique solutions exist (a and c). If there is some error, the range of uncertainty is large if two eigenvectors are nearly parallel (b) while it is not so significant if two eigenvectors are almost orthogonal (d). * represents the range of uncertainty (adopted from Gerald with slight modification).

coincident, there is an infinite number of solutions (non-unique). If the two are parallel but do not coincide, there is no solution. This is the geometrical interpretation of Fredholm alternative.

A small amount of error does not change the range of uncertainty and the result is acceptable (well-posed) or unique (Figure 2.3, d) if two eigenvectors are orthogonal. However, the range of uncertainty increases and shows ill-posedness (Figure 2.3, b) if two eigenvectors are almost parallel. In this case, we may say that the solution is not unique. Large condition number means that the solution may be deflected by errors or noise in the data due to the ill-posed forward model. More specifically, one of the eigenvalues is nearly zero. If we consider the roundoff error in a computer, there is little difference between the non-uniqueness (identifiability) and ill-posedness (instability). Regularization is the term for a various techniques used to restore the well-posedness of an inversion (Koch, 1992). There are some ways to regularize the ill-posedness with minor penalty of the exactness (unbias) and will be discussed later in section 3.

In VSB model, if $f \gg f_0$, then

$$(2-20) \quad g(m) = |v(f)|$$

$$\begin{aligned}
&= \Psi_{\infty} \left(\frac{4\pi^2 f^2}{\alpha^2 r^2} + \frac{1}{r^4} \right)^{1/2} \left\{ \frac{\left(1 + \frac{(1+2B)^2 f^2}{f_o^2} \right)^{1/2}}{\left(1 + \frac{f^2}{f_o^2} \right)^{3/2}} \right\} e^{-\pi f t / Q} \\
&= \left[\Psi_{\infty} f_o^2 (1+2B) \right] \left(\frac{4\pi^2 f^2}{\alpha^2 r^2} + \frac{1}{r^4} \right)^{1/2} \frac{1}{f^2} e^{-\pi f t / Q}
\end{aligned}$$

Since Ψ_{∞} , corner frequency and overshoot are nearly dependent on one another if $f \gg f_o$, each term cannot be resolved uniquely because of the noise and numerical error whose variations are generally greater than the limit of accuracy. When logarithms are taken in the above approximation,

$$\begin{aligned}
\log(g(m)) &= \log(|v_{ff}(f)|) \\
&\approx \log(\Psi_{\infty} f_o^2 (1+2B)) - 2\log(f) \\
&\quad - \frac{\pi t}{Q} e^{\log(f)} + \text{path contribution}
\end{aligned}$$

The high frequency slope beyond the second corner frequency (due to $e^{\log(f)}$) is dependent on Q , source parameters and path contributions. At high frequency, there is an identifiability problem between RDP, overshoot and corner frequency.

For the case where $f < f_o$, we can neglect the attenuation term since $\pi f t / Q$ is very small, making the total attenuation term negligible. The velocity spectrum can be expressed by

$$|v(f)| \approx \Psi_{\infty} \left(\frac{4\pi^2 f^2}{\alpha^2 r^2} + \frac{1}{r^4} \right)^{1/2} \frac{\left[1 + (1+2B)^2 \frac{f^2}{F_o^2} \right]^{1/2}}{\left[1 + \frac{f^2}{F_o^2} \right]^{3/2}}$$

If we take the logarithm in both sides and neglect the near-field term,

$$(2-21) \quad \log |v_{ff}(f)| \approx \log(\Psi_{\infty}) + \frac{1}{2} \log \left((1+2B)^2 \frac{f^2}{F_o^2} + 1 \right) \\ - \frac{3}{2} \log \left(\frac{f^2}{F_o^2} + 1 \right) + \text{path contribution}$$

In this case, Ψ_{∞} , B and F_o are independent of one another. It is to be noted that Ψ_{∞} is determined when $f=0$ after a path contribution correction. There is a possibility of interdependence, and thus trade-off, between overshoot and corner frequency mainly due to the second term in the right hand side of the above equation. However, the third term which is characterized only by the corner frequency reduces the interdependence between overshoot and corner frequency.

In practical application, Ψ_{∞} and overshoot show a great degree of trade-off due to the limited low frequency data below the corner frequency. High frequency approximation illustrated the identifiability problem in parameter estimation previously. It is to be noted that the trade-off

between Ψ_{∞} and overshoot is attributed to the limited bandwidth of data. This type of trade-off induced by the ill-posedness can be removed by obtaining broad-band data. Broad-band data are important not only because Ψ_{∞} and overshoot are mainly constrained at lower frequencies but also because it maintains independence in these parameters.

A similar discussion follows for the HH model. If the frequency is much greater than the corner frequency, then the far-field velocity spectrum (Equation 2-18) approaches

$$(2-22) \quad |v_{ff}(f)| \approx \frac{2\pi f}{\alpha \cdot r} \frac{\Psi_{\infty} F_o^3 (1+6B)}{f^3} \exp\left[-\frac{2\pi f t}{Q}\right]$$

If the logarithm is taken of both sides of the equation, the intercept is determined by $\Psi_{\infty} F_o^3 (1+6B)$ illustrating that the source properties are dependent upon one another again. The above approximation also shows that the HH model will be less well-behaved due to the cube of corner frequency and the larger multiplicative value (6) in B. Small change of either corner frequency or overshoot induces a larger amount of change in corresponding parameters than in the case of the VSB model. As a reference, the VSB model is proportional to the square of the corner frequency and two times the overshoot (Equation 2-20). This result implies that the

inversion process for the HH model would be less well-posed and less stable than the VSB model.

Prewhitening

Let's assume the data satisfy a non-linear theoretical model with random errors. It can be expressed by

$$(2-23) \quad d_{\text{obs}} = g(m) + \varepsilon \quad \text{or}$$

$$\varepsilon = d_{\text{obs}} - g(m)$$

where d_{obs} : data observed
 $g(m)$: theoretical model
 ε : error

For the non-linear equation whose non-linearity is small, Newton's iterative method can be used to find the solution. This is obtained by minimizing the error(ε) in the l_2 norm(least square) sense. If one assumes that the error term(ε) is Gaussian random, one can find m which minimizes the sum of the square of an error.

$$\frac{\partial S}{\partial m} = \frac{\partial \{ (d_{\text{obs}} - g(m))^2 \}}{\partial m} = -2 \frac{\partial g(m)}{\partial m} (d_{\text{obs}} - g(m)) = 0$$

where the misfit function is $S = \{d_{\text{obs}} - g(m)\}^2$.

A matrix expression to find the solution of the above equation by a numerical method(Newton's method) is

$$(2-24) \quad m_{n+1} = m_n + [G^T G]^{-1} (d_{\text{obs}} - g(m_n))$$

where $G = \frac{\partial g(m)}{\partial m}$ and n denotes iteration number. The superscripts T and -1 denote the transpose and inverse matrices respectively. This is the first scheme used in inversion in the next chapter.

The error in the seismic spectrum is the combination of two kinds. The background noise added in the seismogram during recording can be assumed random with a Gaussian distribution. The theoretical error, which is the result of simplification of the source and the path model, can be assumed to be random without serious risk (Tarantola, 1987), although there may be some argument about this assumption. When we apply the above equation to seismic observations in the frequency domain, each data does not have equal variance. In the displacement spectrum, the error at low frequency has much larger variances than that at high frequency. This may lead to problems since the minimum of the sum of the square of all errors is determined mainly by the variances at low frequencies. It means that the data at low frequency have much higher weights than those at high frequency. Equal weight for unbiased inversion can be achieved by prewhitening the data.

If we assume the prewhitening function $w(m)$ is known,

$$(2-25) \quad \frac{d_{obs}}{w(m)} = \frac{g(m)}{w(m)} + \mu$$

where $\mu = \frac{\varepsilon}{w(m)}$.

The new transformed error μ should be random with equal variances for a proper function $w(m)$. When the prewhitening function $w(m)$ is the same as the forward model $g(m)$, the residual(μ) would be most likely random with equal weight variances. The prewhitening process normalizes the variance of the theoretical error, but it distorts the background noise. The effect of distortion by the prewhitening can be reduced by maintaining high signal-to-noise ratios in the frequency domain.

When $w(m)$ is the same as $g(m)$, the misfit function(or least square function) S is expressed by

$$(2-26) \quad S = \sum_{i=1}^n |\mu_i|^2$$

$$= \sum_{i=1}^n \left[\frac{d_i}{g_i} - 1 \right]^2$$

where n is the number of data point and $|\cdot|$ denotes $l-2$ norm.

•This result can be expressed in a matrix notation as

$$(2-27) \quad S = \sum_{i=1}^n \left[\frac{d_i}{g_i} - 1 \right]^2 = \mu^T \cdot \mu = [g_{nn}^{-1} \mathbf{d}_{n1} - \mathbf{I}_{nn} \mathbf{u}_{n1}]^T [g_{nn}^{-1} \mathbf{d}_{n1} - \mathbf{I}_{nn} \mathbf{u}_{n1}]$$

where

$$\mathbf{g}_{nn} = \begin{bmatrix} g_1(m) & . & . & 0 \\ 0 & g_2(m) & . & . \\ . & . & . & . \\ 0 & . & . & g_n(m) \end{bmatrix}$$

$$\mathbf{d}_{n1} = \begin{bmatrix} d_{obs1} \\ . \\ . \\ d_{obsn} \end{bmatrix}$$

$$\mathbf{I}_{nn} = \begin{bmatrix} 1 & . & . & 0 \\ . & 1 & . & . \\ . & . & . & . \\ 0 & . & . & 1 \end{bmatrix} \quad \text{and}$$

$$\mathbf{U}_{n1} = \begin{bmatrix} 1 \\ 1 \\ . \\ 1 \end{bmatrix}$$

The subscript represents the size of the matrix. Therefore, since \mathbf{g}_{nn} is a diagonal matrix,

$$\begin{aligned} s &= \mathbf{d}_{n1}^T \mathbf{g}_{nn}^{-1} \mathbf{g}_{nn}^{-1} \mathbf{d}_{n1} - \mathbf{d}_{n1}^T \mathbf{g}_{nn}^{-1} \mathbf{I}_{nn} \mathbf{U}_{n1} \\ &\quad - \mathbf{U}_{n1}^T \mathbf{I}_{nn} \mathbf{g}_{nn} \mathbf{d}_{n1} + \mathbf{U}_{n1}^T \mathbf{I}_{nn} \mathbf{U}_{n1} \\ &= \mathbf{d}_{n1}^T (\mathbf{g}_{nn} \mathbf{g}_{nn})^{-1} \mathbf{d}_{n1} - \mathbf{d}_{n1}^T (\mathbf{g}_{nn} \mathbf{g}_{nn})^{-1} \mathbf{g}_{nn} \mathbf{U}_{n1} \\ &\quad - \mathbf{U}_{n1}^T \mathbf{g}_{nn}^T (\mathbf{g}_{nn} \mathbf{g}_{nn})^{-1} + \mathbf{U}_{n1}^T (\mathbf{g}_{nn} \mathbf{g}_{nn})^{-1} \mathbf{g}_{nn} \mathbf{U}_{n1} \\ &= (\mathbf{d}_{n1} - \mathbf{g}_{nn} \mathbf{U}_{n1})^T (\mathbf{g}_{nn} \mathbf{g}_{nn})^{-1} (\mathbf{d}_{n1} - \mathbf{g}_{nn} \mathbf{U}_{n1}) \end{aligned}$$

$$= (\mathbf{d}_{n1} - \mathbf{g}_{n1}(m))^T \mathbf{C}_{nn}^{-1} (\mathbf{d}_{n1} - \mathbf{g}_{n1}(m))$$

where $\mathbf{C}_{nn} = \begin{bmatrix} g_1^2(m) & . & . & 0 \\ . & g_2^2(m) & . & . \\ . & . & . & . \\ 0 & . & . & g_n^2(m) \end{bmatrix}$

$$\mathbf{g}_{n1} = \begin{bmatrix} g_1(m) \\ . \\ . \\ g_n(m) \end{bmatrix}$$

Now, for minimal S,

$$(2-28) \quad \frac{\partial S}{\partial m} = 0 = \mathbf{G}^T \mathbf{C}^{-1} (\mathbf{d} - \mathbf{g}(m))$$

where $\mathbf{G} = \frac{\partial \mathbf{g}_{n1}}{\partial m}$

If Newton's iterative method is applied,

$$(2-29) \quad m_{n+1} = m_n - \left[\frac{\partial^2 S}{\partial m^2} \right]_{m_n}^{-1} \left[\frac{\partial S}{\partial m} \right]_{m_n}$$

$$= m_n - \left[\mathbf{G}_n^T \mathbf{C}^{-1} \mathbf{G}_n - \left(\frac{\partial^2}{\partial m^2} \right)_{m_n} - (d_{\text{obs}} - \mathbf{g}(m_n)) \right]^{-1} \left[\mathbf{G}_n^T \mathbf{C}^{-1} (d_{\text{obs}} - \mathbf{g}(m_n)) \right]$$

where n denotes the number of iteration and $\frac{\partial^2 S}{\partial m^2}$ denotes Hessian. The above implementation can be simplified if we assume that the second derivative of g is small compared to the first derivative, G.

$$\begin{aligned}
(2-30) \quad m_{n+1} &= m_n - \left[\frac{\partial^2 S}{\partial m^2} \right]_{m_n}^{-1} \left[\frac{\partial S}{\partial m} \right]_{m_n} \\
&\approx m_n - [G_n^T C^{-1} G_n]^{-1} [G_n^T C^{-1} (d_{\text{obs}} - g(m_n))]
\end{aligned}$$

The relaxation parameter ρ ($0 < \rho \leq 1$) can be applied to the second term of right hand side of the above equation to make the iteration convergent. Equation 2-30 becomes

$$(2-31) \quad m_{n+1} = m_n - \rho [G_n^T C^{-1} G_n]^{-1} [G_n^T C^{-1} (d_{\text{obs}} - g(m_n))] \quad 0 < \rho \leq 1$$

This is the prewhitening process without damping in a non-linear inversion. The new matrix C which results from the prewhitening process can be interpreted as the data covariance matrix (Tarantola, 1987) whose variance at specific data points is proportional to the square of its own value or the weight matrix (Wiggins, 1972) whose diagonal elements are the square of inverse of their own values.

Application to the Forward Model

There are several ways to reduce the condition number and improve the inversion process. The first is the normalization of parameters by transformation (Bates and Watts, 1988) from a priori information. Since it is clear that the parameters in the forward model (Equation 2-13 through 2-16) are all positive and the order of the parameters can be determined by a few trials, each parameter

can be scaled so the derivative matrices have nearly equal order. A simple normalization of Equations 2-13 and 2-14 can be expressed as

$$(2-32) \quad |u(f)| = k_1 a \left(\frac{1}{\alpha r} + \frac{1}{2\pi f r} \right)^{1/2} \left\{ \frac{\left(1 + \frac{k_2^2 b^2 f^2}{k_3^2 c^2} \right)^{1/2}}{\left(1 + \frac{f^2}{k_3^2 c^2} \right)^{3/2}} \right\} e^{-\pi f t / (k_4 d)}$$

$$(2-33) \quad |v(f)| = k_1 a \left(\frac{4\pi^2 f^2}{\alpha^2 r^2} + \frac{1}{r^4} \right)^{1/2} \left\{ \frac{\left(1 + \frac{k_2^2 b^2 f^2}{k_3^2 c^2} \right)^{1/2}}{\left(1 + \frac{f^2}{k_3^2 c^2} \right)^{3/2}} \right\} e^{-\pi f t / (k_4 d)}$$

with the unknowns of a , b , c , and d

$$a = \frac{\Psi_\infty}{k_1} \quad b = \frac{1+2B}{k_2}$$

$$c = \frac{F_0}{k_3} \quad \text{and} \quad d = \frac{Q}{k_4}$$

Determination of k_i ($i=1,2,3,4$) can be done either from a priori information or by the automatic scaling during iteration. While a fixed normalization factor from a priori information is simple and consistent throughout the whole process of inversion, an incorrect initial guess seriously affects the rate of convergence. On the other hand, an

The second procedure designed to reduce the condition number is known as the Marquardt-Levenberg (ML) method. This technique is also called Levenberg-Marquardt (LM) method, damped least squares, or ridge regression (Lines and Treitel, 1984). Not only can it reduce the condition number without any significant bias in model parameters but its rate of convergence is quite fast when properly applied. This technique is so powerful that it is widely used in ill-posed geophysical inversion. The algorithm for the ML method is

$$(2-36) \quad m_{n+1} = m_n - (G_n^T G_n + kA)^{-1} G_n^T (g - \eta) \quad 0 < \rho \leq 1$$

where

- G : derivative matrix of misfit function S with respect to model parameter (Frechet matrix)
- k : damping parameter
- A : an identity matrix or the diagonal matrix of Hessian
- g : theoretical model
- η : observed data

Comparing this procedure to the prewhitening process without damping (Equation 2-31), this algorithm includes damping. The above equation can be derived in a straight forward manner by applying a Lagrangian multiplier to the non-linear problem and constraining the variance of model parameter to be a constant (Marquardt, 1963). The damped

diagonal matrix makes the inversion stable by adding a D.C. offset to the Hessian matrix which may have singular eigenvalues. Therefore, the ML method always has a solution and its solution is unique (Lines and Treitel, 1984; Koch, 1992) in small residual problems since the inverse of diagonal matrix (simplified Hessian) is also a diagonal matrix whose components are the reciprocals of the diagonal components of the original matrix. When the residual is large, the second derivative of Hessian matrix is dominant and the existence of solution cannot be guaranteed in the ML method since the Hessian may be neither a diagonal matrix nor a diagonal dominant matrix. When the relaxation parameter is applied to Equation 2-36, it becomes

$$(2-37) \quad m_{n+1} = m_n - \rho(G_n^T C^{-1} G_n + kA)^{-1} [G_n^T C^{-1} (d_{\text{obs}} - g)]$$

This scheme includes a damping factor. There is no limitation on the value of k , it can range from zero to infinity. When k is zero, this scheme is the same as Gauss-Newton scheme (Equation 2.31). If k is large, the Hessian behaves almost like an identity matrix due to the dominance of the damping parameter. Therefore, this is the same as the steepest descent method. This technique is not introduced or tested here because of its well known slow rate of convergence. An intermediate damping parameter provides both regularization and speed of the convergence. The k used here

is different from the corner frequency k in the Haskell type forward models.

The same problem can be solved using the method suggested by Tarantola(1987; Tarantola and Valette, 1982 (a) and (b)). He derived an a posteriori probability density in the model space by introducing the statistics of the data covariance matrix and model covariance matrix. The data covariance matrix contains the uncertainty of each data set while the model covariance matrix contains the uncertainty of the model parameters. Each element of the data or model covariance matrix is composed of the covariance of each element of the data or model parameters. The best-fit estimates can be obtained by maximizing the a posteriori probability or minimizing the misfit function. The prewhitened data inversion scheme is equivalent to the Tarantola's scheme with data covariance matrix whose diagonal elements are g_i^2 ($i=1,2,\dots,n$) and off-diagonal elements are all zero as shown previously. In this case, the model covariance matrix plays the role of reducing the condition number. Diagonal matrices of data or model covariance imply independence between the elements of the data or model parameters. The equation is written here with slight modification of the a priori model parameter m_{pr} to m_{n-1} for actual implementation.

$$(2-38) \quad m_{n+1} = m_n + \rho$$

$$\cdot \{[G_n^T C_d^{-1} G_n + C_m^{-1}]^{-1} [G_n^T C_d^{-1} (g(m_n) - d_{obs}) + C_m^{-1} (m_n - m_{n-1})]\}$$

where $G_n = \left[\frac{\partial g}{\partial m} \right]_{m=m_n}$

As the iteration goes on, if the model converges to the data, $(m_n - m_{n-1})$ will approach zero. Therefore, the term $C_m^{-1} (m_n - m_{n-1})$ can be neglected after a few iterations. The inverse of the model covariance matrix (C_m^{-1}) in the first bracket of Equation 2-38 plays the role of the diagonal matrix A in the ML method (Equation 2-37) that was used for regularization. The convergence path is equivalent to that of the ML method with prewhitening. This scheme is used for both prewhitening and damping.

Convergence

Contrary to the linear case, non-linear least square inversion does not guarantee the convergence of the parameter by iteration. In an iterative inversion, three sequences of convergence should be checked.

- 1) $m_{n+1} - m_n \rightarrow 0$ (Convergence of parameter estimates)
- 2) $g(m_{n+1}) - g(m_n) \rightarrow 0$ (Convergence of function estimates)
- 3) $G(m_{n+1}) - G(m_n) \rightarrow 0$ (Convergence of gradient estimates)

as iteration goes to infinity. Unfortunately, one type of convergence does not guarantee others. If any one of these convergence criteria is not satisfied, the estimation by the

iterative non-linear inversion is not stable. Convergence in the non-linear least square inversion can be obtained by the introduction of a relaxation parameter. This term is also called a damping factor in some literature. I use the term relaxation parameter, as introduced in this paper, to avoid confusion with the damping parameter which is used to correct the ill-posedness and reduce the trade-offs. The relaxation parameter should lie between zero and one. Even though there are some ways to calculate the maximum possible value of the relaxation parameter or radius of trust region (Seber and Wild, 1989), it is reasonable to fix the parameter as a constant. Acceptable values for the relaxation parameter can be obtained by a trial and error experiment.

Termination Criteria

For the criteria of iteration termination, either of the above mentioned convergence criteria can be used. If all of the above estimates go to zero, iteration can be terminated safely. Unfortunately, however, they do not go to zero even with ideal data due to roundoff error in the computer. It is necessary, therefore, to obtain sufficiently small values beyond which the accuracy of the estimates does not improve significantly.

From the basic assumption of the least square method, the derivative of misfit function equals to zero for minimization and the gradient convergence would be the best

criteria for termination. This requirement, however, is too strict to estimate the parameters in many cases (Seber and Wild, 1989). Instead, the termination of iteration can be done from the rate of convergence of parameters. If the rate of change of each parameter is sufficiently small, we can safely terminate the iteration. In multivariate inversion, each parameter may have different rates of change.

Therefore, the sum of the square of convergence rates is used as a criteria in this work. Along with the termination of the iteration, each parameter can be fixed if the rate of change of that specific parameter is very small. It reduces the calculational load since the fixed value can be safely eliminated from the calculation in the next inversion step. There are a few other methods used to calculate the criteria of iteration termination (Seber and Wild, 1989) which guarantee convergence. But the simplest way to determine the criteria of iteration termination is the trial and error method with data whose parameters are already known. Synthetic data are a good example. Trial and error methods with synthetic data will be used to determine when the iteration can be stopped safely in this work.

Simultaneous Inversion

Simultaneous inversion for the source parameters from the several data sets can be done with slight modification. Several different spaces can be operated at the same time by

allowing each subsystem(Lanczos, 1961) to be independent in a single matrix. The independency between the subsystems(or data set) can be maintained by introducing appropriate zero matrices. An example problem is given using the VSB model and sources with different overshoot(B);

$$\begin{array}{c}
 \left[\begin{array}{cc|c|c}
 g_{11}^1 & g_{12}^1 & g_{13}^1 & 0 \\
 \vdots & \vdots & \vdots & \vdots \\
 g_{m_1 1}^1 & g_{m_1 2}^1 & g_{m_1 3}^1 & 0 \\
 \hline
 g_{11}^2 & g_{12}^2 & 0 & g_{13}^2 \\
 \vdots & \vdots & \vdots & \vdots \\
 g_{m_2 1}^2 & g_{m_2 2}^2 & 0 & g_{m_2 3}^2
 \end{array} \right] \begin{array}{c} \text{RDP} \\ f_0 \\ B^1 \\ B^2 \end{array} = \begin{array}{c} d_1^1 \\ \vdots \\ d_{m_1 1}^1 \\ \hline d_1^2 \\ \vdots \\ d_{m_2 1}^2 \end{array} \\
 (m_1+m_2) \times 4 \qquad 4 \times 1 \qquad (m_1+m_2) \times 1
 \end{array}$$

The representation can be used for the simultaneous inversion, where g_{jk}^i is the i -th coefficient at j -th frequency for k -th parameter and d_j^i is the i -th data at j -th frequency. Each subsystem is divided by a dotted line. Since this method shares common information about RDP and f_0 and discriminates different overshoots from each data set independently, the estimated parameters may be less biased from unknown effects such as detailed path structure, anisotropic propagation, and heterogeneous scattering which generally result in source parameter fluctuation when such data is inverted. Figure 2.4 is the result of simultaneous inversion from two sets of ideal data with different overshoots. The estimated parameters are matched to the

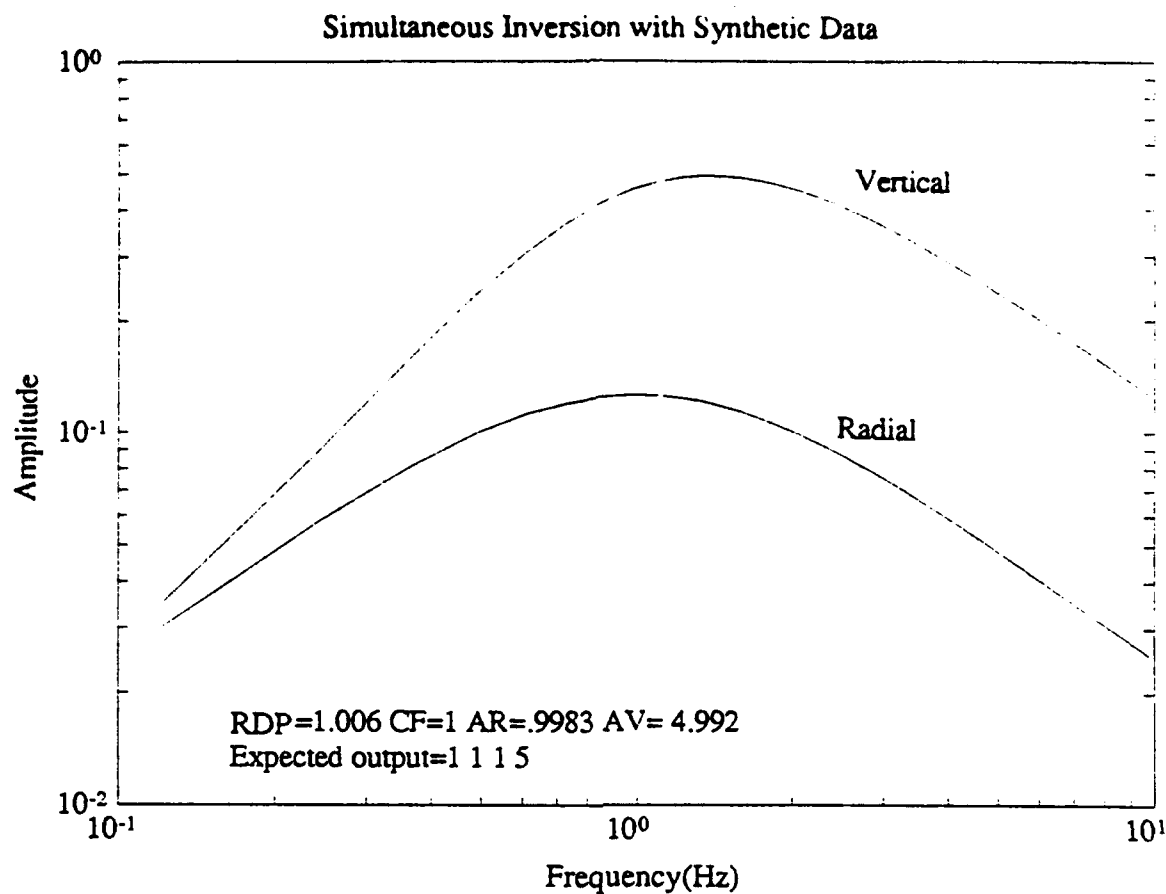


Figure 2.3 Simultaneous inversion for the same Ψ_∞ and corner frequency with different overshoot. The data were generated synthetically. Estimated parameters ($\Psi_\infty=1.006$, $F_0=1$, $B_{\text{radial}}=-0.0$ ($A_{\text{radial}}=0.9983$), $B_{\text{vertical}}=1.996$ ($A_{\text{vertical}}=4.992$)) are close to the expected values ($\Psi_\infty=1$, $F_0=1$, $B_{\text{radial}}=0$ ($A_{\text{radial}}=1$), $B_{\text{vertical}}=2$ ($A_{\text{vertical}}=5$)).

expected values. The same argument can be done for the same source parameters with different Q's. Generally, it can be formulated by a general form as

$$(2-39) \quad \begin{bmatrix} C_1 & | & U_1 & | & 0 & | & 0 \\ \hline & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ C_n & | & 0 & | & 0 & | & U_n \\ \hline & & & & & & \end{bmatrix} \begin{bmatrix} X_c \\ \hline X_{p1} \\ \vdots \\ X_{pn} \end{bmatrix} = \begin{bmatrix} D_1 \\ \hline \vdots \\ D_n \end{bmatrix}$$

where C_i : coefficient of common factor for i-th data
 U_i : coefficient dependent on the i-th data set
 X_c : common parameters
 X_{pi} : specific parameters dependent only on the i-th data set
and D_i : i-th data set.

The main disadvantage of the simultaneous inversion is that it requires a huge amount of computer capacity since the coefficient and data matrices are generally large.

Other Considerations in Inversion

The interdependence(correlation) between the data, which seems to be mainly due to the simplified forward model and

partly due to the windowing effect in data, may be taken into account in an inversion by an autoregressive model (Lines and Treitel, 1984). This can be done by introducing non-zero off-diagonal elements to the Frechet matrix (Bates and Watts, 1988). If the data can be expressed by

$$\bar{d} = g(m) + \xi_1(|d_1 - d_{i-1}|) + \xi_2(|d_1 - d_{i-2}|) + \dots + \xi_k(|d_1 - d_{i-k}|)$$

where ξ_k is the linear correlation between the elements of the data which should be removed for the independence of parameters. For an ideal case of randomly distributed residual, all ξ_k should be zero. The i -th off-diagonal term in the data covariance matrix (C) corresponds to ξ_i in an inversion (Box and Jenkins, 1970). However, this methodology was not adopted in this work because of the ineffectiveness of computation and the small improvement in bias of the parameters, even though it may be a better way in the statistical sense of unbiasedness and the basic assumption of randomness of the residual (Bates and Watts, 1988).

The inverse process can be simplified also by introducing conditional linearity (Bates and Watts, 1988). An a (corresponding to Ψ_∞) in Equations 2-32 through 2-35 is conditionally linear because the derivative of the forward model with respect to a does not involve an a . Thus, this parameter can be treated as a constant in the non-linear inversion and determined later by linear inversion. This

method does not reduce the condition number directly, but it gives a better chance for stability simply because the number of parameters to be determined in the non-linear inversion is reduced. After a few trials with this method, it was decided not to apply it because there was no clear advantage in synthetic source parameter inversion.

Inversion with Bootstrap

One of the most exciting recent development in statistics may be the bootstrap concept introduced by Efron(1979). Bootstrapping is a computer-based, robust and useful method to estimate parameters. Furthermore, its theory is simple and application is wide. Bootstrap method is found to be especially powerful for the treatment of limited data set. This property of the bootstrap method may be important for the source parameter inversion since the steady state RDP estimation is mostly affected by a few data points in the low frequencies. Furthermore, low frequency bands are easily contaminated by the numerical and background noises which makes it hard to estimate parameters reliably by a single estimation.

Efron(1979 and 1981) summarized the theoretical background in his papers and Efron and Tibshirani(1986) summarized the applicability of the bootstrap method. McLaughlin(1988) and Koch(1992) have used this method in

measuring uncertainty in magnitude estimation and in solving S wave structure and Poisson's ratio in Germany respectively.

Since the thorough explanation of the theory and the applicability are documented in the previously mentioned papers, the discussion here will be brief and restricted.

Let's assume a given random variable X with unknown distribution F .

$$X_1, X_2, \dots, X_n \sim \text{iid } F$$

Each data set shows independent and identical distribution(iid) which can be assumed without serious concerns in most observed data. F is an unknown distribution. Since we don't know the distribution F , we cannot estimate the true first and second central moment(eg, mean and standard deviation). Instead, if we assume the empirical probability distribution \hat{F} where each X_i has the same probability in \hat{F} . Then,

$$(2-39) \quad \sigma(\hat{F}) = \left[\frac{\mu_2(\hat{F})}{n} \right]^{1/2}$$

$$\mu_2(\hat{F}) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

where $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

These are the bootstrap estimates. We can calculate the bootstrap estimates numerically as many times as we want, by the Monte Carlo algorithm. The procedure for bootstrap with Monte Carlo algorithm is composed of three steps as is illustrated in Figure 2.4; (i) draw bootstrap samples using a random number generator with uniform distribution. This means that random number should be generated randomly with replacement. The best sample size of bootstrap is the same size as the original data size (Efron and Tibshirani, 1986). (ii) for each bootstrap sample, estimate statistics. (iii) calculate statistics from the bootstrap statistics. For example, unbiased bootstrap mean and standard deviation can be estimated by

$$(2-40) \quad \hat{\theta}^*(\bullet) = \frac{\sum_{b=1}^B \theta^*(b)}{B}$$

$$\hat{\sigma}_B = \left(\frac{\sum_{b=1}^B \{\hat{\theta}^*(b) - \hat{\theta}^*(\bullet)\}^2}{B-1} \right)^{1/2}$$

where
$$\hat{\theta}^*(b) = \frac{\sum_{i=1}^n (x_i - \bar{x})}{n}$$

and B = total number of bootstrap process.

If B goes to infinity, then the mean and the standard error will approach to the true mean and standard deviation in an assumed bootstrap distribution \hat{F} by the Central Limit

Theorem. Even though there still remains the problem how close the bootstrap estimates are to the true statistics, Efron and Tibshirani(1986) claimed that bootstrap is a useful tool to estimate statistics by comparing with other methods(Table 2 in their paper).

Contrary to the basic assumption of normality of the data in the least square non-linear source parameter inversion, the non-linear algorithm with bootstrapping can estimate the source parameters regardless of the distribution of the residual. It releases the possible restrictions on the application of least square inversion such as randomly and normally behaved path effect and outlier-free residuals. The procedure takes into account outliers caused by unexpected effects like low frequency noise impact on source parameter estimates using the least square method.

The bootstrap procedure implemented by the Monte Carlo method is relatively simple. First, sample N number of data randomly with replacement from N data elements. Next, estimate the source parameters by the non-linear least square inversion. This procedure is done several times. Last, we can accept the bootstrap estimates of source parameters if the distribution of the sampled parameters is normal. If some of the data cannot explained by the theoretical forward model and behave like outliers, the distribution of the bootstrap results will not be normally distributed(shown in the next chapter with some empirical test and theoretical

considerations). The shape of the distribution of the bootstrap statistics is a good indicator that the inversion process has been done properly. When the shape is far from bell-shape, it tells that the least square estimate is not proper for this type of data and the estimates are not reliable and generally behave as outliers when compared others showing Gaussian. Even in those cases, however, the optimal estimates may be selected by the characteristics of the shape of the distribution which corresponds or is close to the l_1 norm-like estimates. This property needs more theoretical consideration and is not used in estimating source parameters in this paper.

There is no clear settlement about the number of iterations necessary to obtain reliable estimates. McLaughlin(1988) has chosen a minimum of ten times of bootstrap for the estimation of simple mean while Koch(1992) claimed at least 100 calculations for a reasonable value for a standard deviation and more than 1000 times for the estimation of confidence interval based on the work by Efron and Tibshirani(1986).

Summary

Several explosion source models were investigated. Generalized Haskell type of model is simple to formulate for the source parameter inversion. This model can be applied to various kinds of high frequency decay models without

reformulating the whole equation. Generalized equation has significant advantages in programming and in parameterizing high-frequency roll-off compared to the individual formula representing each source model. As an attenuation model, simple frequency independent Q was used. Basic properties of the forward model, such as existence and uniqueness of the solution, stability and regularization in an inversion process are reviewed. This model is non-linear with respect to the corner frequency, overshoot and Q . It also shows trade-offs between parameters at high frequencies (non-uniqueness). Mathematical derivation shows that it is impossible to separate steady state RDP, corner frequency and overshoot effects at high frequencies beyond the corner frequency. It implies that the resolution of individual parameters can be done successfully only with the broad-band data. Various techniques to regularize the forward model are reviewed and applied to the explosion source model.

Various inversion methods and their relations were reviewed. Based on a review of the inversion methods, inversion process by prewhitening was developed. This process is equivalent to Tarantola's technique which uses data and model covariance matrices. Simultaneous inversion for the source parameters from several data set was developed, although this method will not be applied to the observed data. Convergence and termination criteria for an

CHAPTER 3

SYNTHETIC TEST

Before applying inversion method to the observed data, it is necessary to examine that which scheme is the most optimal for the source parameter inversion, what causes the bias of the parameters, and if there is any way to reduce the bias, if any. Synthetic tests are the best way to examine above questions since one knows the expected values.

This chapter is composed of three parts. The first describes a set of synthetic tests to examine different inversion schemes. The data will be generated directly from the source model in the frequency domain and the test will be done in the homogeneous-full space. Since the data are generated from the forward model, it is easy to verify the inversion process such as the prewhitening process, difference between the schemes, noise effect, bias, and trade-offs between parameters. The process of non-linear inversion is not known well even though it is used frequently in the field of geophysics and thus the process of each scheme will be empirically investigated. Several of the schemes discussed earlier will be tested. Each scheme has the advantages and disadvantages and the optimal scheme may

iterative inversion are also reviewed. Total rate of parameter change is chosen as a termination criteria.

Other considerations for the inversion programming such as non-diagonal Hessian matrix by an autoregressive model and partial linearity of steady state RDP in the forward model were done, but they are not applied in the source parameter inversion because of the ineffectiveness of the computation.

The non-parametric bootstrap method was introduced and discussed briefly. The bootstrap procedure was also reviewed. Since this is independent of the distribution of the data, the bootstrap method with a non-linear inversion may be appropriate to resolve the source parameters from uncertain data sets. The applicability of the bootstrap in the non-linear source parameter inversion was discussed. The empirical test for the characterization of bootstrap will be done in Chapter 3.

be changed from model to model. Optimal scheme for the source parameter inversion will be chosen based on these tests. Modified scheme for efficient calculation will be derived and tested if it works well in the source parameter inversion. Parameter biases by the near-field propagation term when the inversion assumes far-field propagation model will be discussed.

The second part describes inversion tests with synthetic seismograms generated from homogeneous half-space and layered structures. It is important for source parameter inversion to simulate the realistic data since the spectra calculated from the synthetic seismograms generally do not match perfectly with the forward model due to the path correction, free-surface interaction, and some numerical noise. These tests are also important to visualize the degree of bias of the parameters by these secondary effects. The bias and their trade-offs by the secondary effects will be tested. The importance of specific wave propagation effects such as surface waves will be studied with these trials. High frequency information is not generally used in the source estimation since it is believed that the higher frequency is more susceptible to the minor features of the structure and since different source model results in different source estimation at this frequencies (Aki, 1976). It is also impossible to express each parameter individually at this frequency bandwidth as shown earlier. Nevertheless, the

spectra show consistent results at high frequencies since the body wave effect is dominant while the low frequencies show the mixed spectra of body waves, surface waves, spall, and the numerical and background noise. The effect of minor features is not so significant in evaluating source parameters in the averaging process to the smooth forward model as the secondary effect. High frequency approximation is important in checking consistency throughout the whole data sets.

Finally, non-parametric bootstrapping will be tested and compared with the least-square inversion based on the simple assumption of normal distribution of the residuals. In source parameter inversion in the spectral domain, any norm based method has weakness in estimating steady state RDP, overshoot and corner frequency. They are easily biased by an introduction of secondary effect and noise at low frequency data points since resolution of each parameter is possible only at this frequency band. Maintaining high signal-to-noise ratio limits the number of available data set and still show fair amount of variances because of the secondary effect. Bootstrap is not only independent of the distribution of the population, robust for an analysis of limited data, but also powerful for an extension of available data set. The empirical tests will show that the bootstrap can extract unbiased information from the data with outliers.

Synthetic Test

Prior to application to observational data, the inversion program was tested with synthetic data. These data were generated from the original model (Equation 2-16) of VSB's. The parameters used were $\Psi_{\infty} = 1.0$, $f_0 = 1.0$, $Q = 50$ and no overshoot ($B=0$) for Brune model and an overshoot ($B=2$) for the VSB model. There is no practical and mathematical difference whether displacement, velocity or acceleration data are used in an inversion if the path effect is corrected since the inverse process is performed from the source spectra. Prewhitening is a part of inversion process in this work rather than prewhitening the data before inversion.

Program tests for the ω^{-3} (Helmberger-Hadley) model were not performed because there is no difference with the ω^{-2} (von Seggern-Blanford) model except the formulation in the forward model. For simplicity, the source-receiver distance was assumed to be 1 km and the compressional velocity was taken as 1 km/sec in the homogeneous full-space. The frequency of the data ranges from 0.05 Hz to 50 Hz with sampling interval of 0.05 Hz. This bandwidth is sufficient to resolve steady state RDP, overshoot, corner frequency and attenuation. Frequency independent attenuation was applied.

Synthetic Data without Overshoot

Tests were performed with (1) ideal data, (2) ideal data with theoretical noise, and (3) ideal data with background noise. The near-field propagation path correction was

neglected for the purposes of the test. Theoretical noise is taken to be uniformly distributed in the entire frequency range with a mean of unity. This noise contribution is multiplied to each of corresponding amplitude point(modulus) in the frequency domain so that the residuals are uniformly distributed if prewhitened. Since multiplication in the amplitude domain is addition in logarithmic space, the expected bias of the parameter by applying the theoretical noise should be zero in an ideal case.

$$\log(\text{mean}(\text{theoretical noise}))=0$$

The actual application of theoretical error to the ideal data, however, may cause bias in the estimation of steady state RDP since only a few low frequency data points are important in determining it.

Background noise in this numerical trials is the same type of random noise labeled theoretical noise except its mean is scaled by -28 dB(5 %) of the maximum amplitude of ideal acceleration spectra at each frequency. The slope of -1 in log-log space was applied to this noise to simulate velocity data(acceleration data integrated once). The flat(in frequency) background noise in the acceleration seismogram will have the slope of -1 after integration to velocity. This noise was added to the ideal data.

Comparison of inversion scheme

Comparison was performed among several different schemes;

(1) inversion with neither prewhitening nor damping (Gauss-Newton method),

(2) inversion with prewhitening,

(3) inversion with damping (ML method), and

(4) inversion with prewhitening and damping.

Functional expressions of above four schemes are as follows;

$$(1) \quad m_{n+1} = m_n - \rho \left(G_n^T G_n \right)^{-1} \left[G_n^T (d_{obs} - g) \right]$$

$$(2) \quad m_{n+1} = m_n - \rho \left(G_n^T C^{-1} G_n \right)^{-1} \left[G_n^T C^{-1} (d_{obs} - g) \right] \quad (2-29 \text{ again})$$

$$(3) \quad m_{n+1} = m_n - (G_n^T G_n + \rho A)^{-1} G_n^T (g - \eta) \quad (2-35 \text{ again})$$

$$(4) \quad m_{n+1} = m_n + \rho \cdot$$

$$\cdot \{ [G_n^T C_d^{-1} G_n + C_m^{-1}]^{-1} [G_n^T C_d^{-1} (g(m_n) - d_{obs}) + C_m^{-1} (m_n - m_{n-1})] \}$$

(2-36 again)

where $G_n = \left[\frac{\partial g}{\partial m} \right]_{m=m_n}$

Several different initial values were input into each scheme to verify that the resulting models were global minima. Initial values are taken randomly. They are distributed within the expected parameters with variances of the same order. Total rate of change of $2.5e-9$ was used as a criterion of iteration truncation. This value is sufficient

to assure that all parameters converge to their own specific values. For a posterior analysis, it is not allowed to set each parameter as a constant during iteration even in the case of a very small rate of change. Several values of the damping parameter (10 - 300) were tested in the damped schemes (Scheme 3 and 4).

The estimated parameters are the same as the expected values through all of the schemes when the data were generated directly from the Brune's model ($\Psi_{\infty}=1$, $B=0$, $F_0=1.5$, and $Q=50$). The outputs were not strongly affected by 30 different initial values (Figure 3.1 a,b, and c). There are a few cases where the estimates represent a local minima. Note that logarithmic steady state RDPs were taken in histogram to show the distribution distinctively.

Figure 3.2 shows the number of iterations taken to get the assigned termination criteria for various initial inputs. Both the simple Gauss-Newton method (Scheme 1) and the prewhitening scheme (Scheme 2) converge rapidly to the expected values, generally less than 20 iterations. The numbers in the middle of each plot are average numbers of iterations from 30 different trials. The median was used as a statistical average of iterations since the distribution of iteration number is far from a normal distribution. When the damped inversion was investigated without prewhitening (Scheme 3), its convergence rate was quite slow and required many more iterations to arrive at the same criteria of

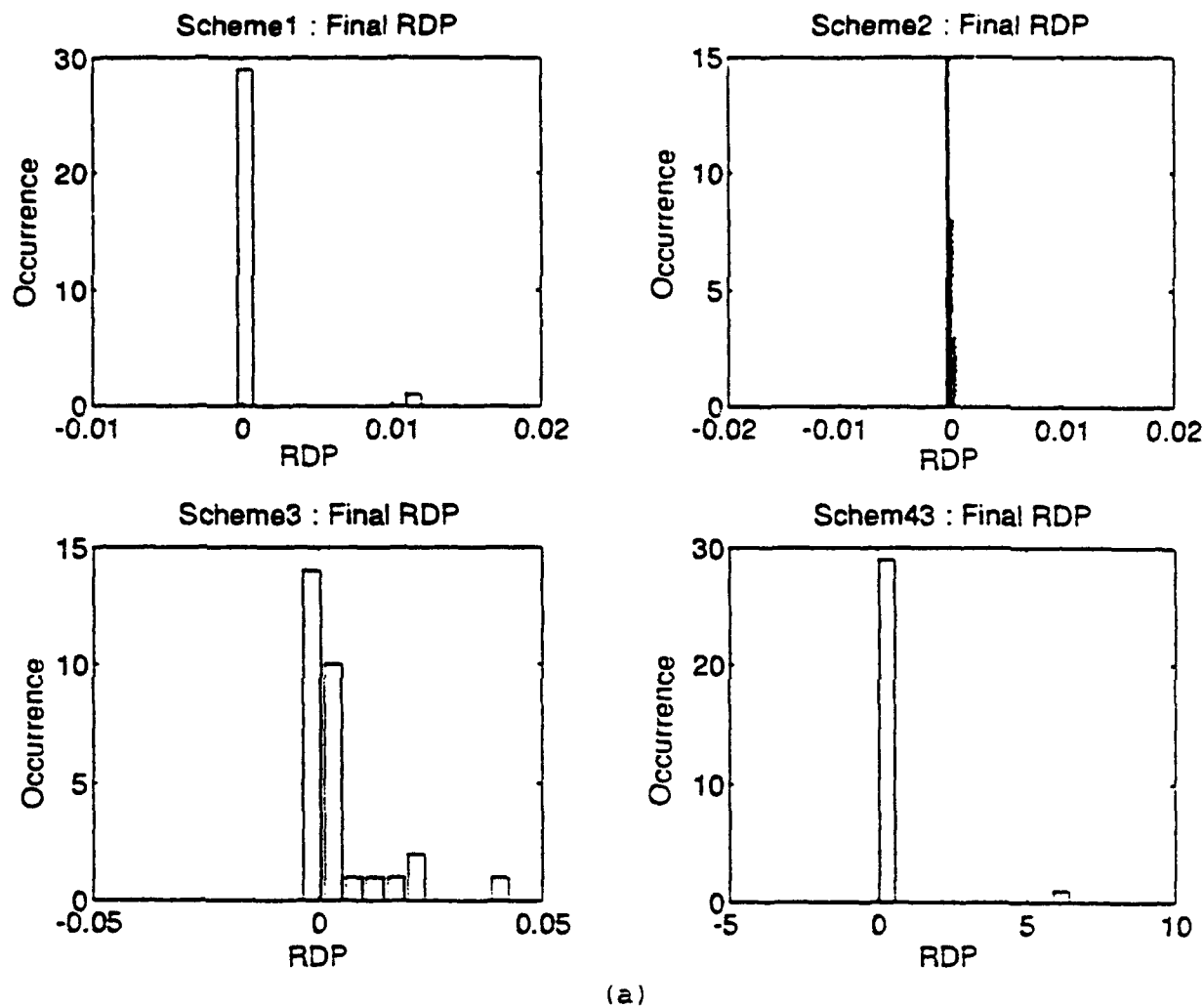


Figure 3.1 Estimated parameters from 30 different initial values. Histograms show that all schemes resolve the parameters well regardless of given initial values when the ideal data are used. (a) Ψ_{∞} ; (b) corner frequency; (c) Q .

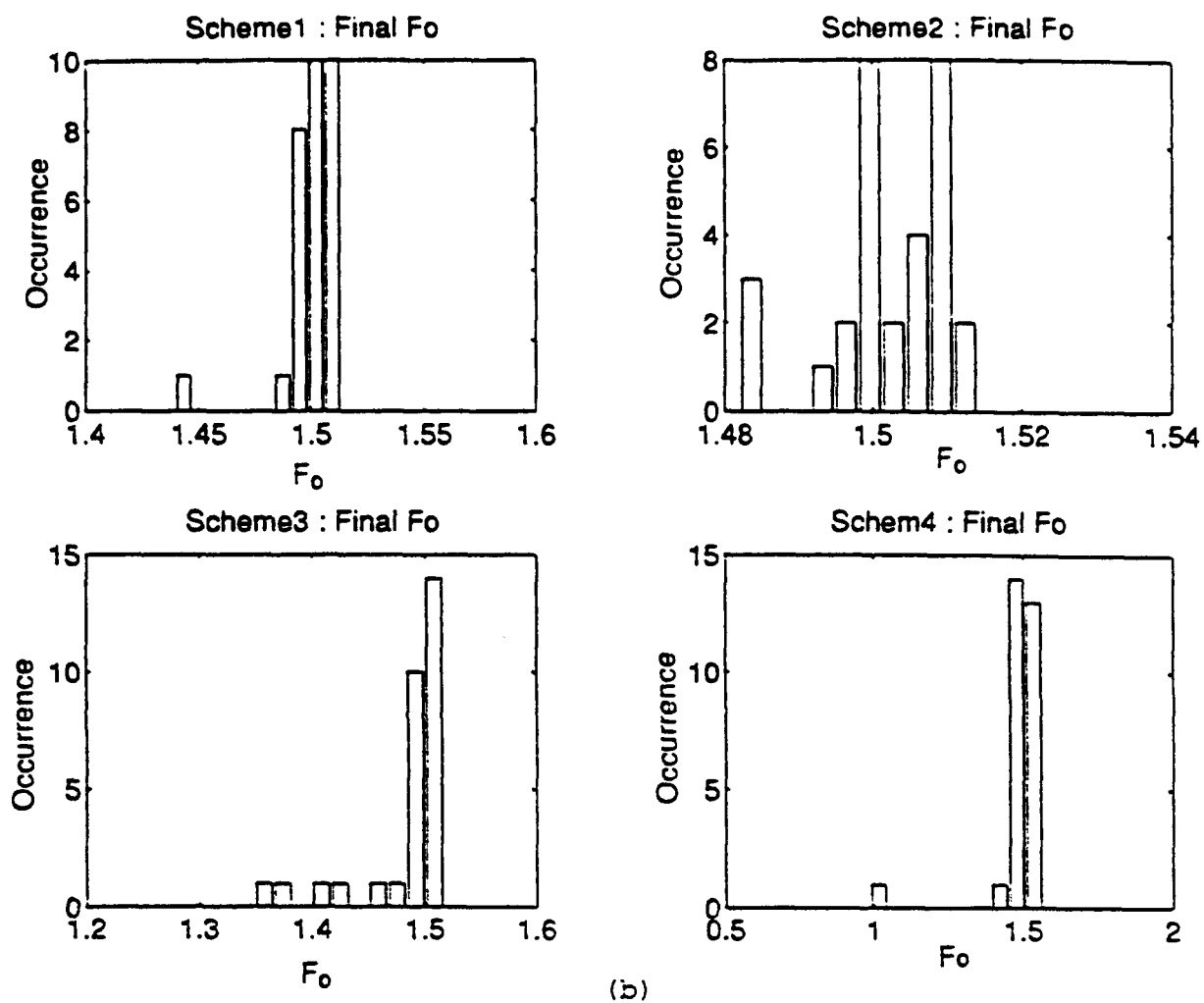
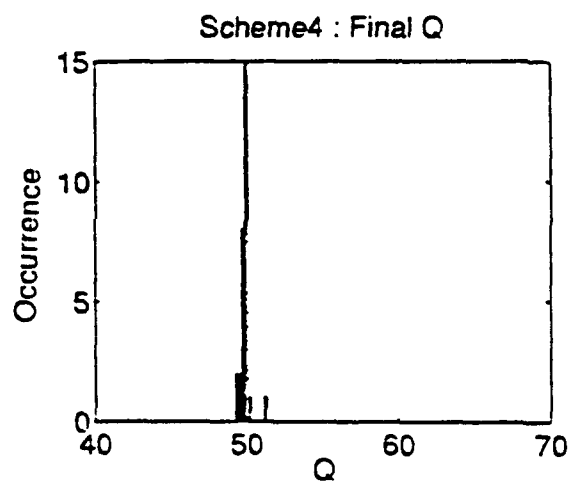
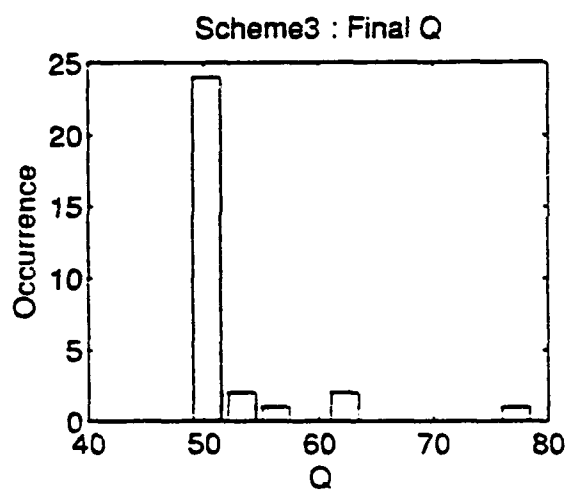
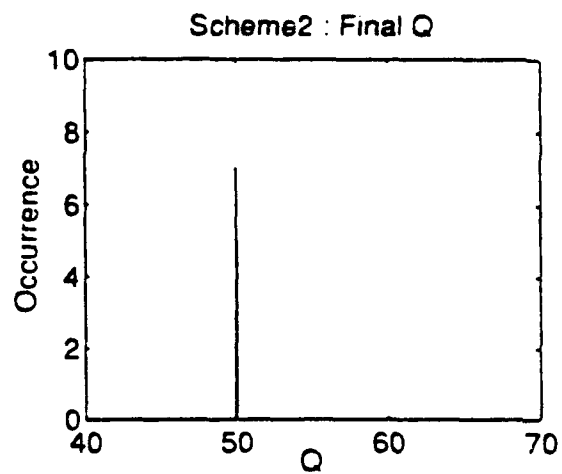
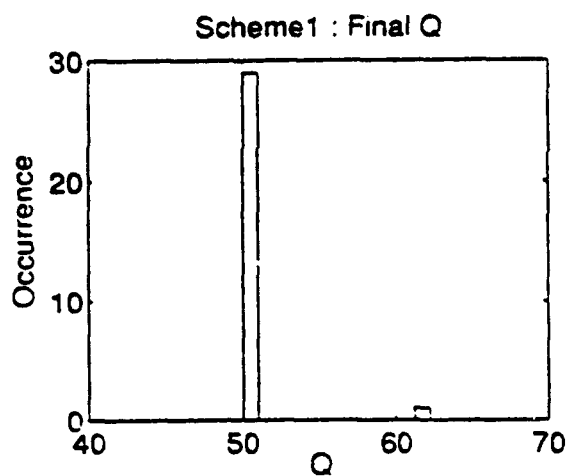


Figure 3.1 Continued.



(c)

Figure 3.1 Continued.

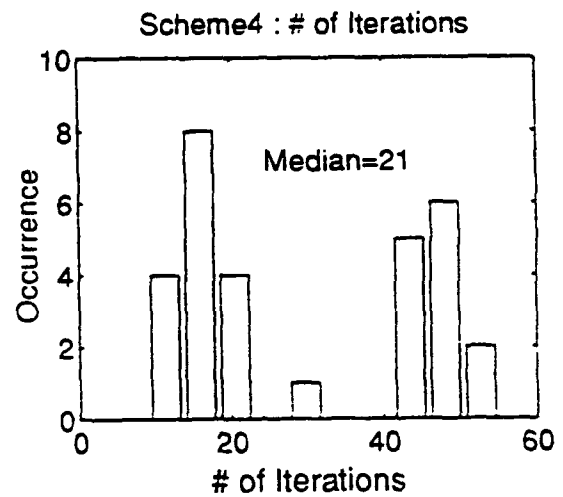
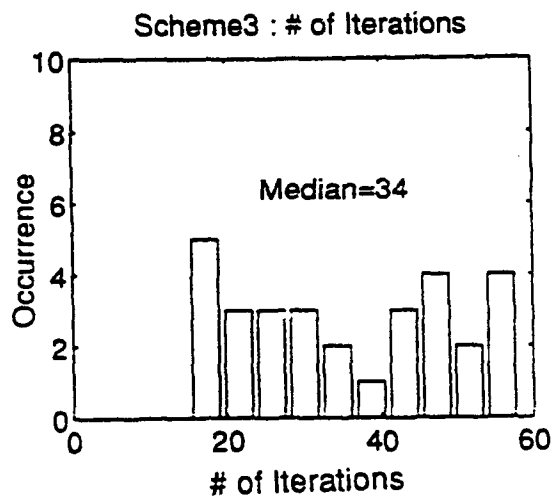
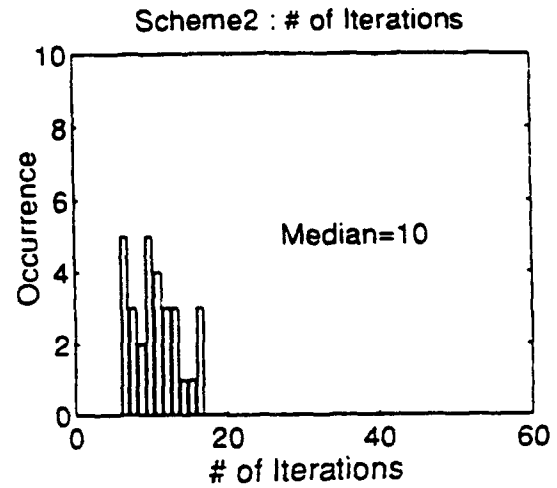
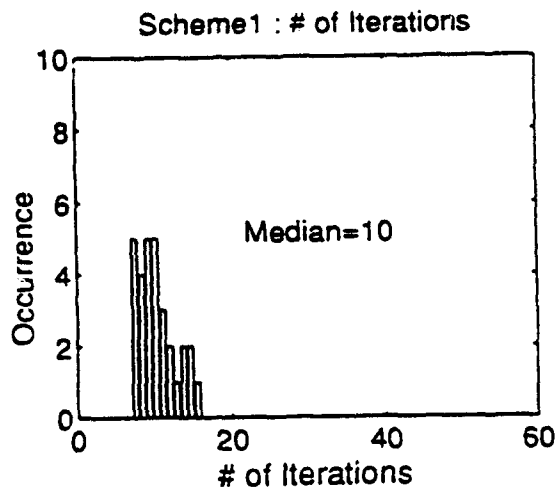


Figure 3.2 The number of iterations taken to get an assigned termination criteria(2.5×10^{-9}) at each scheme. Scheme 3 shows the slowest convergence rate.

termination of iteration due to the de-emphasis of the high frequencies and the value of the damping parameter. The scheme with prewhitening and damping (Scheme 4) shows two peaks in its histogram. This scheme converges either quickly or quite slowly. It converges as quickly as Scheme 2 when the damping parameter plays a minor role in the eigenvalues and converges as slowly as Scheme 3 when the damping parameter is important.

Correlation coefficients (Appendix C) were calculated from the covariance matrices to investigate the linear dependence between the parameters. Correlation coefficient can be defined as follows;

$$\rho_{i,j} = \frac{C_{i,j}}{\sqrt{C_{i,i}}\sqrt{C_{j,j}}}$$

where $\rho_{i,j}$: Correlation coefficient between parameters
 $C_{i,j}$: Elements of covariance matrix

It is the normalized correlation between the parameters, generically related to the forward model. High correlation coefficient (nearly ± 1) means that there are linear relationship between two parameters. Thus, these two parameters can trade-off easily if a small amount of noise is added.

Figure 3.3(a,b, and c) displays the correlation coefficient at each iteration from 30 different calculations

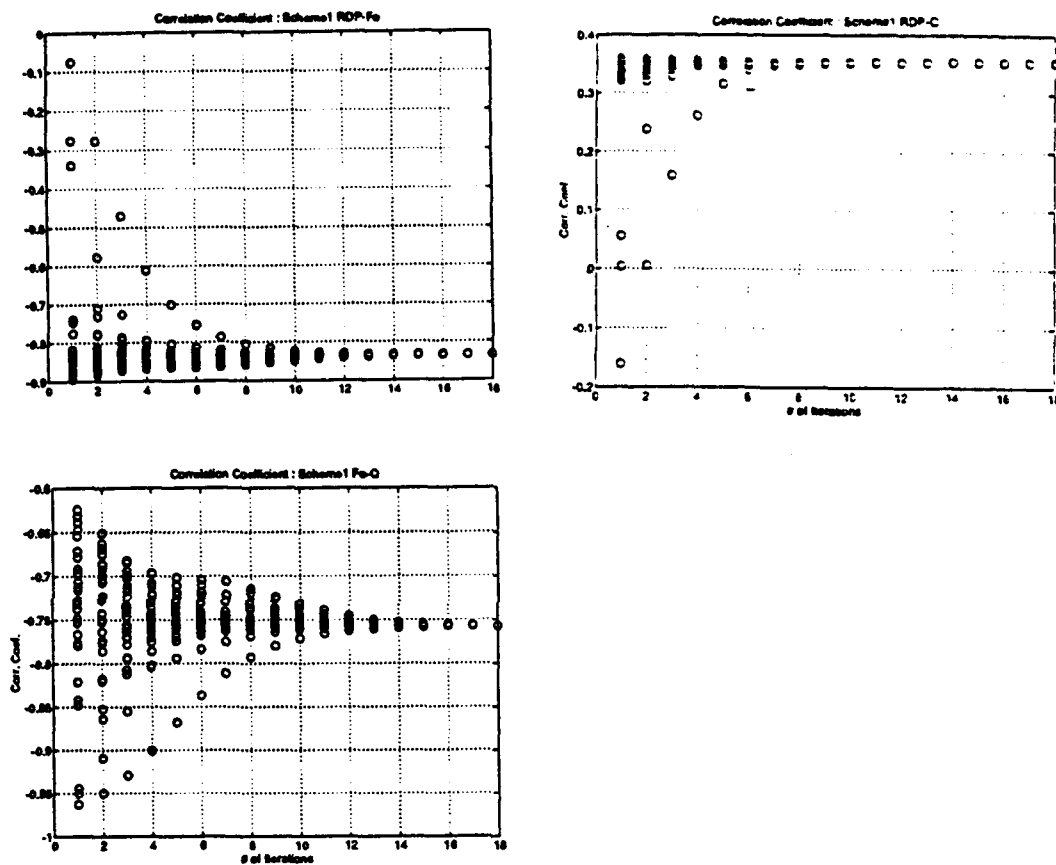


Figure 3.3 Correlation coefficients between parameters with Scheme 1. 30 correlation coefficients (a) between Ψ_{∞} and f_0 , (b) Ψ_{∞} and Q and (c) f_0 and Q are plotted. Correlation coefficients between Ψ_{∞} and f_0 is large.

for Scheme 1. They illustrate how the correlation coefficient changes as the iteration proceeds. Open circles denote the correlation coefficient at each iteration. The closed circles denote the final correlation coefficient at each inversion. Correlation coefficients converge to a specific values as the iteration proceeds. Convergence of correlation coefficient implies the iteration has proceeded sufficiently enough to arrive at the minima before it meets the termination criterion. Despite high correlation coefficients between Ψ_{∞} and corner frequency and between corner frequency and Q in this scheme, no trade-off between parameters occurs since exact data was used. Correlation coefficient between Ψ_{∞} and Q is low as can be expected from the forward model.

Scheme 2 shows high correlation between Ψ_{∞} and corner frequency (Figure 3.4,a). It is worse than that of Scheme 1. This scheme, however, displays reduced correlation coefficients with Q (Figure 3.4,b and c), dropping from 0.36 to 0.22 between Ψ_{∞} and Q and from -0.76 to -0.46 between corner frequency and Q . Prewhitening is necessary to resolve corner frequency and Q independently.

Scheme 3 with a damping parameter of 10 leads to no improvement in the correlation coefficient between steady state RDP and corner frequency (Figure 3.5,a). The correlation coefficient between Ψ_{∞} and Q (Figure 3.5,b) is 0.14, which means that there is no serious trade-off between

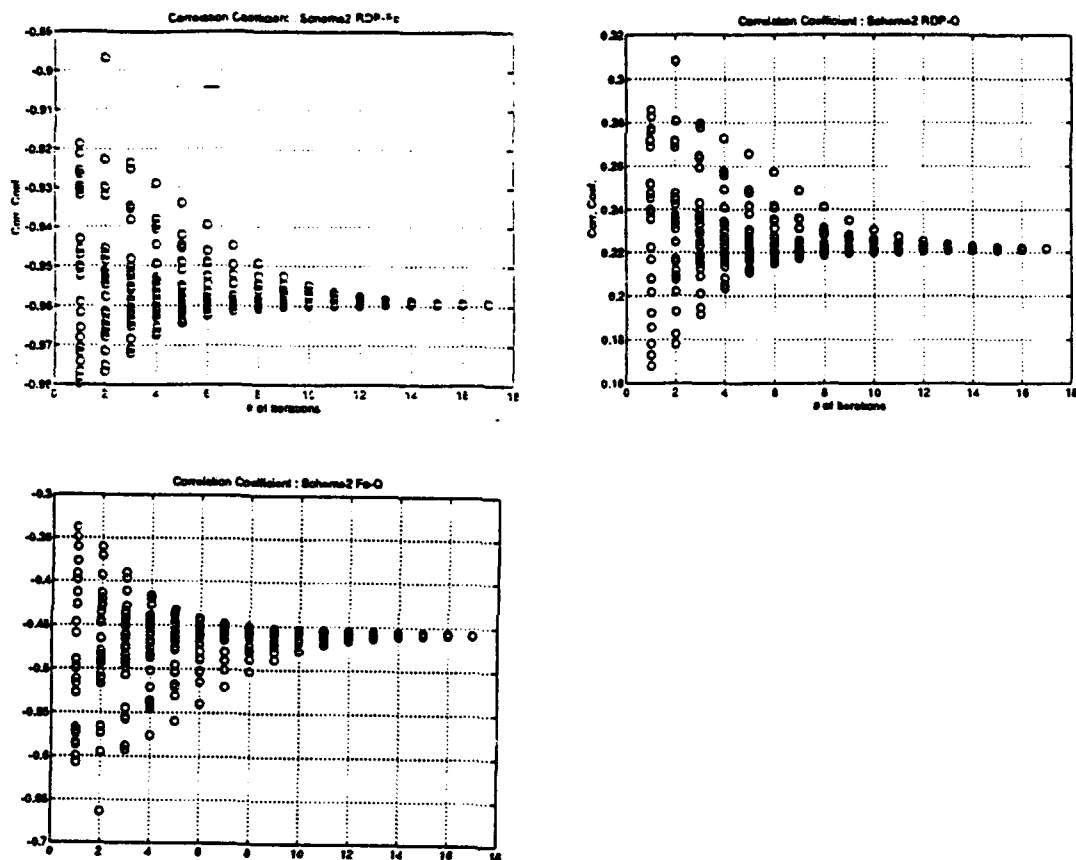


Figure 3.4 Correlation coefficients between parameters with Scheme 2. 30 correlation coefficients (a) between Ψ_{∞} and f_0 , (b) Ψ_{∞} and Q and (c) f_0 and Q are plotted. Correlation coefficients between Ψ_{∞} and f_0 is large.

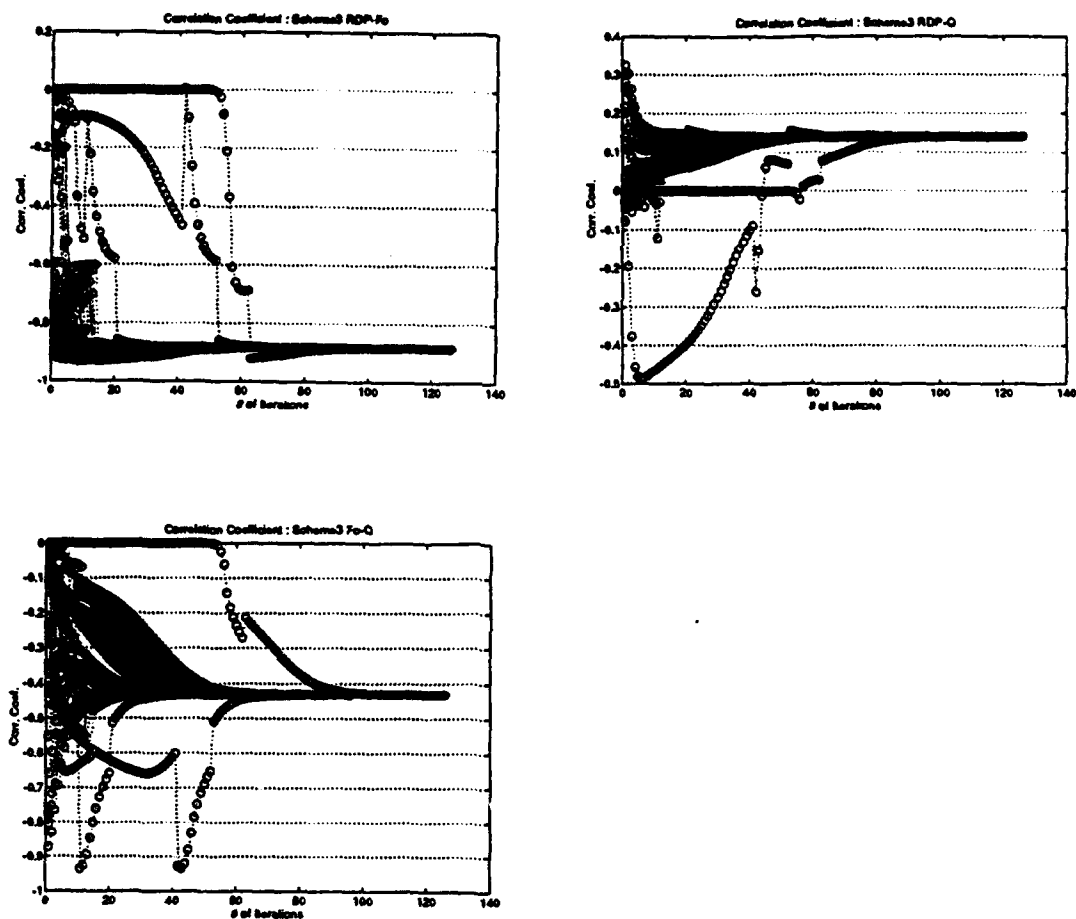


Figure 3.5 Correlation coefficients between parameters when Scheme 3 is used. 30 correlation coefficients (a) between Ψ_{∞} and f_0 , (b) Ψ_{∞} and Q , and (c) f_0 and Q are plotted as circles.

these parameters. The correlation coefficient between corner frequency and Q (Figure 3.5,c) also drops to -0.42 even though prewhitening is not applied in this scheme. When the damping parameter is increased from 10 to 100, the correlation coefficient between steady state RDP and corner frequency tends to be smaller (Figure 3.6). Correlation coefficient was checked at every 10 iterations in this case. There is a general trend for the correlation coefficient and the condition number to decrease as the damping parameter increases. Large damping factor, however, leads to a slower convergence rate (Figure 3.7). One hundred iterations with a small damping parameter (=10) is increased to 654 iterations with a large damping parameter (100) for the same criteria of iteration termination. If a damping parameter goes to infinity, Scheme 3 takes the path of the steepest descent and shows very slow convergence rate. For consistency, the same initial inputs and relaxation parameters were used. The biases of the estimates due to different damping parameters are very small (Figure 3.8).

When the damping parameter is applied in an inversion, there can be abrupt discontinuities in the correlation coefficient between iterations (Figure 3.5). The jump of correlation coefficient results from the change of automatic normalization factor. Figure 3.9 plots the first 17 iterations from Scheme 3 with damping parameter of 10. As shown in this figure, the normalization factor for steady

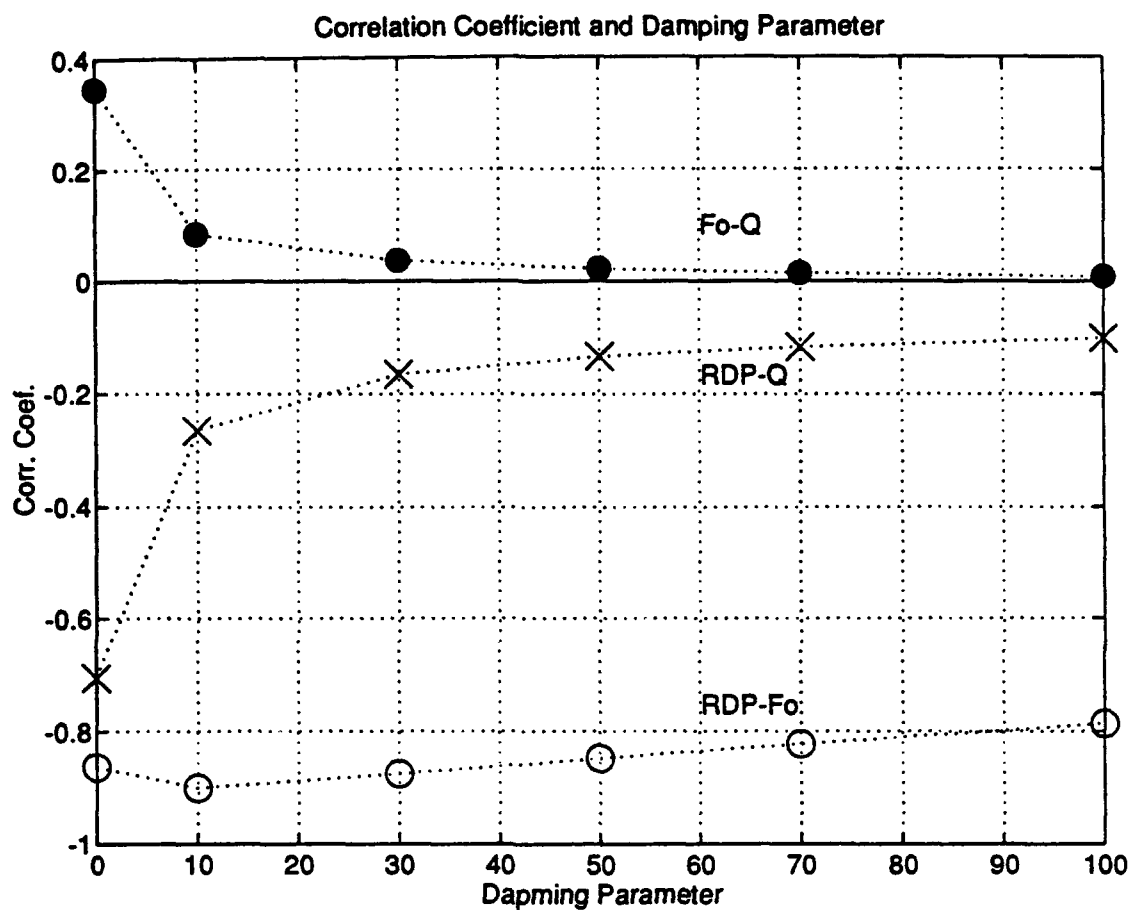


Figure 3.6 Correlation coefficient and damping parameter. There is a trend that the correlation coefficient decreases as the damping parameter increases.

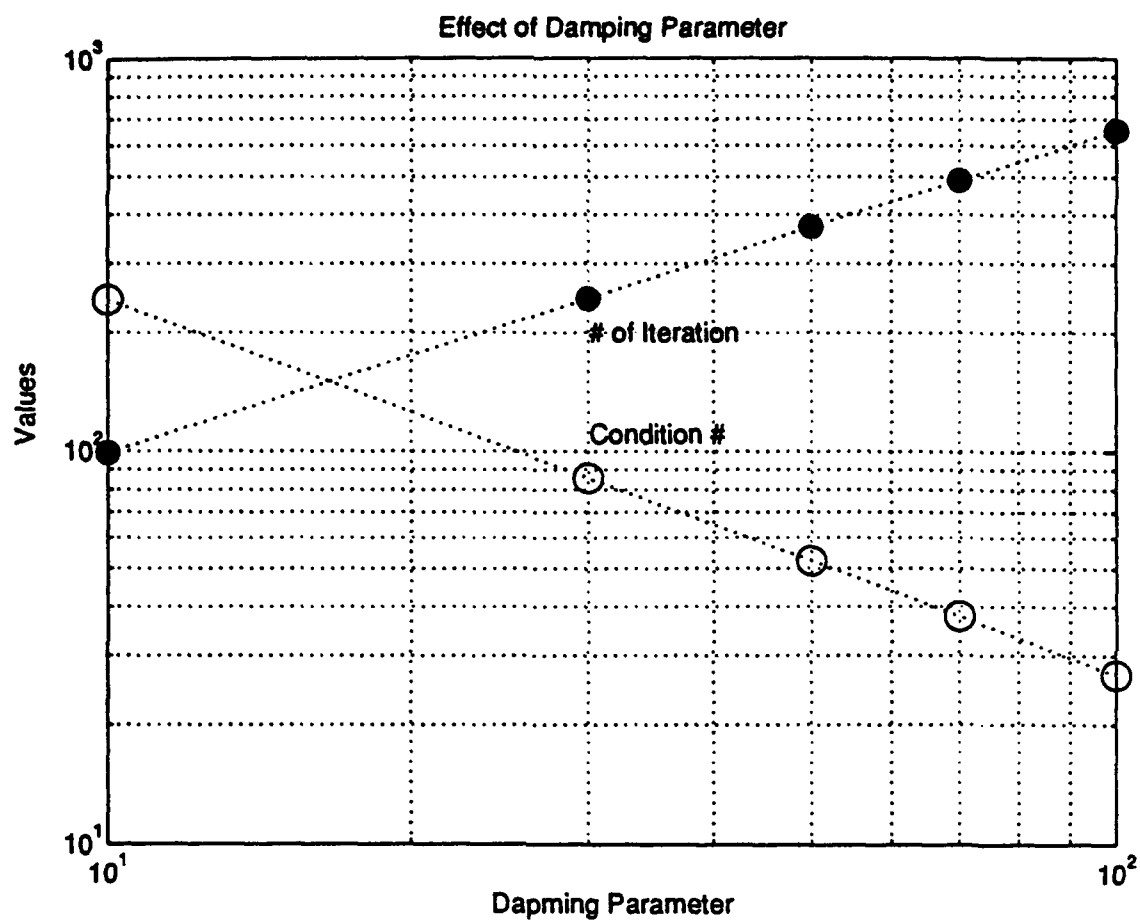


Figure 3.7 Relation of damping parameter, number of iteration and condition number. Condition number decreases while the total number of iteration increases as the damping parameter increases.

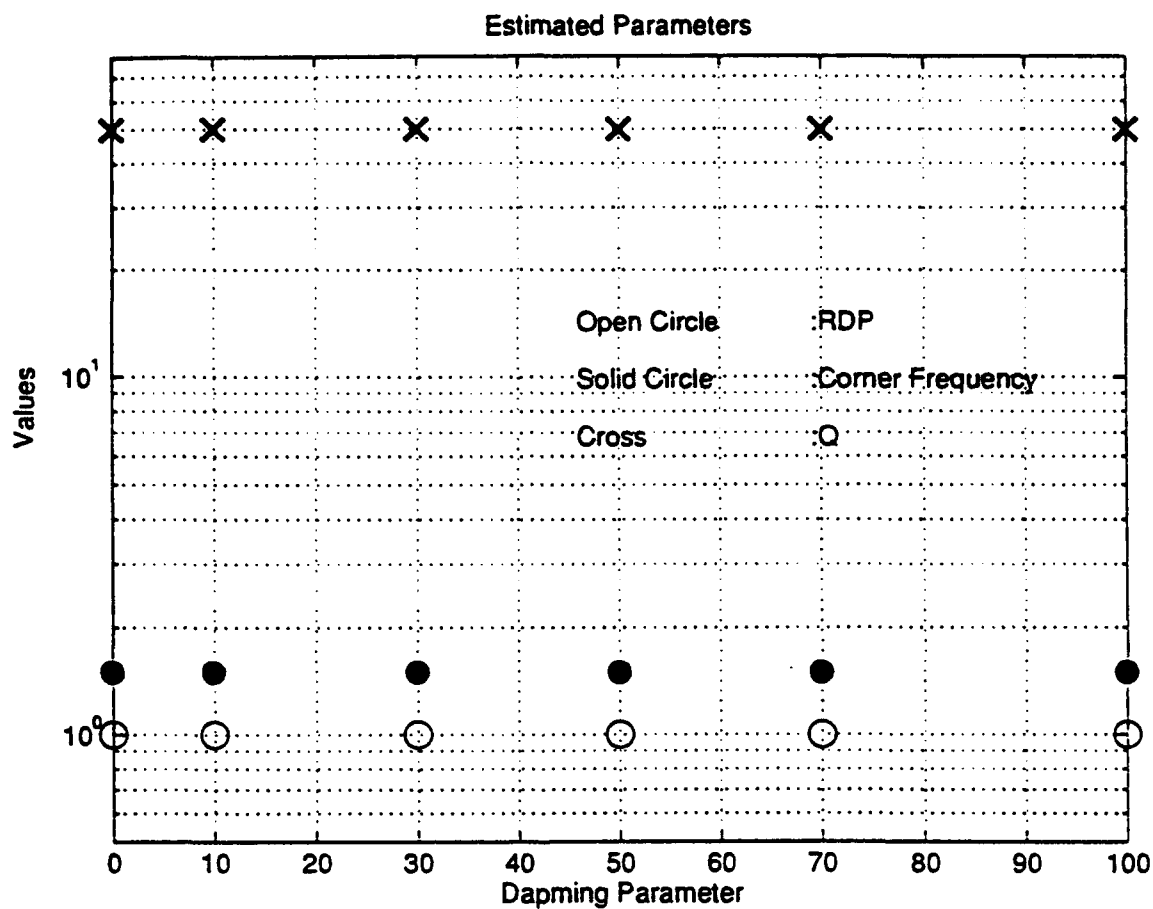


Figure 3.8 There are little bias of the parameters at given values of dapming parameter less than 100. Ψ_{∞} 's are denoted as open circles, corner frequencies are solid circles and Q's are crosses.

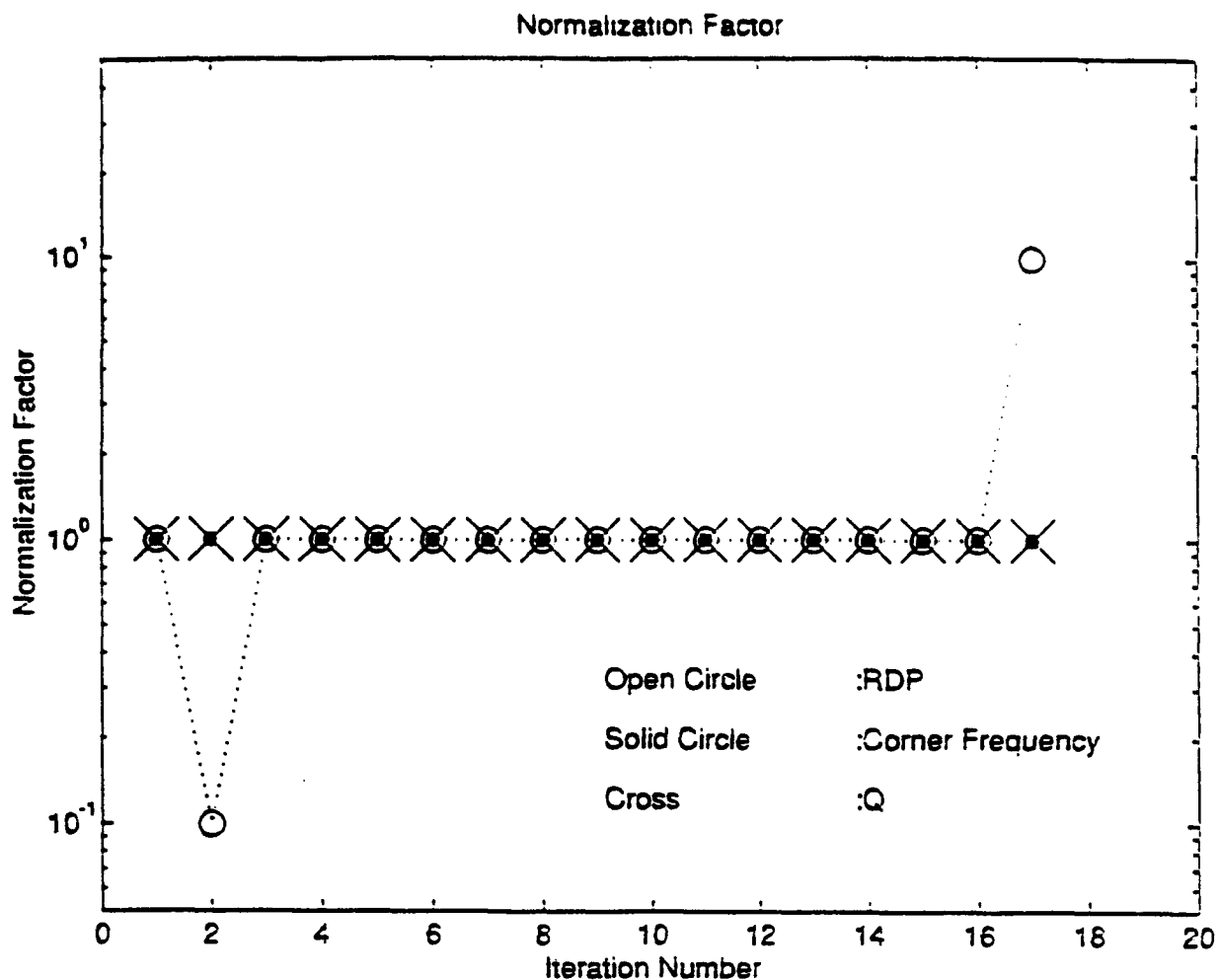


Figure 3.9 This figure illustrates why the jump of correlation coefficient occurred in Scheme 3. Ψ_{∞} was rescaled at the second and 17th iteration. Open circles denote Ψ_{∞} 's, solid circles are corner frequencies, and crosses are Q's.

state RDP changes at the second and 17th iteration which correspond to abrupt changes of the correlation coefficients and the condition numbers displayed in Figure 3.10. With the change of normalization factor, there is a corresponding change in the significance of the damping parameter. The change of the weight of damping parameter results in a discontinuous behavior in the correlation coefficient and condition number. This behavior does not happen in Scheme 1 and 2 (no damping) and in Scheme 3 with large damping parameter because the large damping parameter is dominant compared to the effect of normalization.

Scheme 4 (Figure 3.11, a, b, and c) shows the mixed characteristic of Scheme 2 and Scheme 3 in its correlation coefficient. Correlation coefficient between STEADY STATE RDP and corner frequency becomes worse than that of Scheme 3 just as in the case of Scheme 1 and Scheme 2. However, it is not clear whether the prewhitening is the main reason for this deteriorating correlation coefficient. There is slight improvement in its correlation coefficient at each parameter when they are compared to the corresponding values in Scheme 2. Correlation coefficients between parameters and characteristics from the different inversions are summarized in Table 3.1.

Figure 3.12 summarizes the effect of the choice of initial model parameters in the convergence or divergence of the iterations. Open circles denote the initial values from

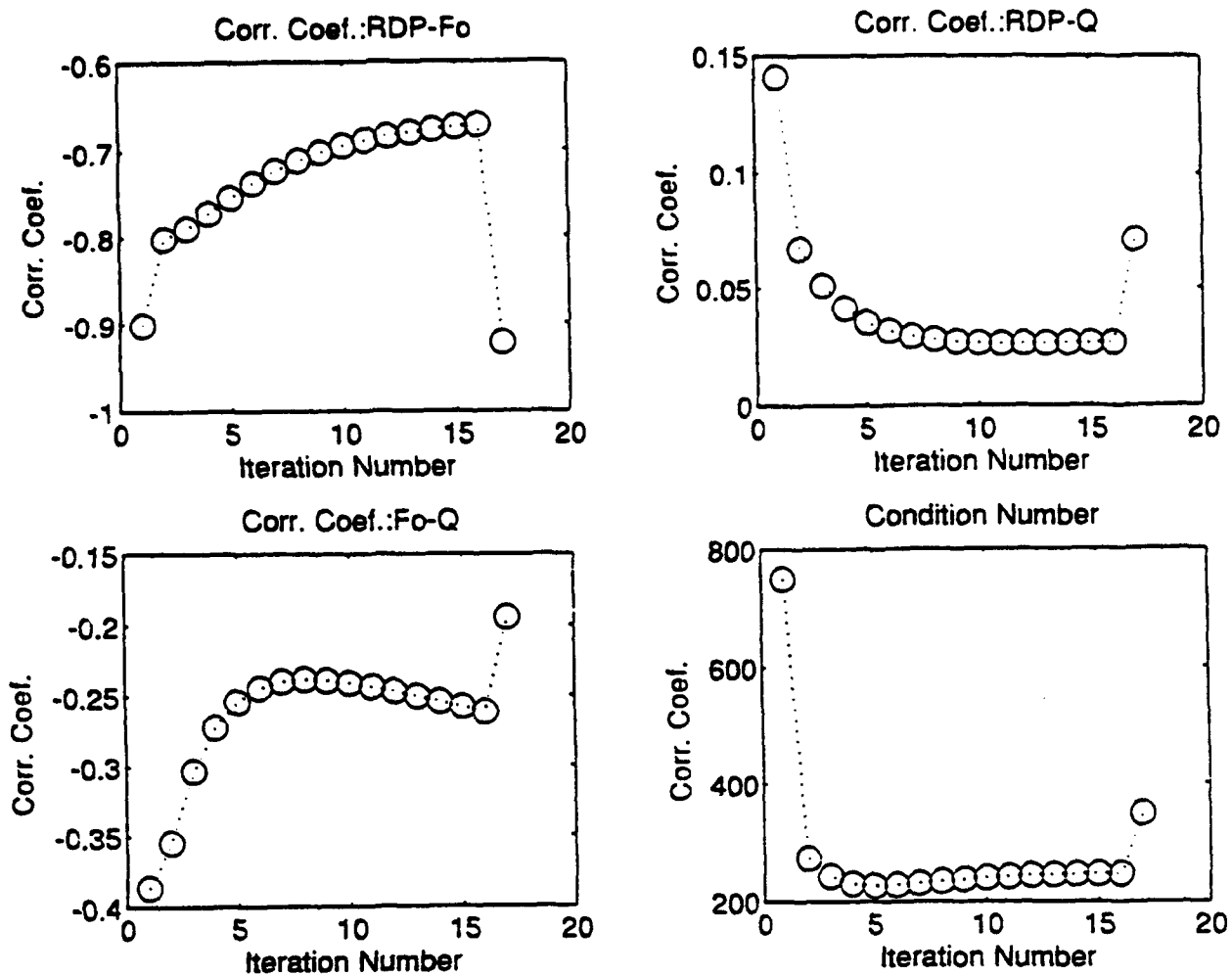


Figure 3.10 The change of correlation coefficients and condition numbers. The jumps occurred at the second and the 17th iterations.

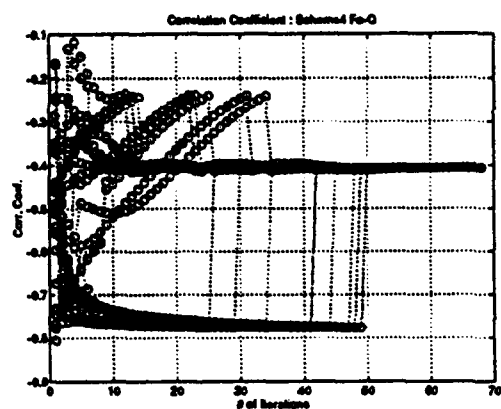
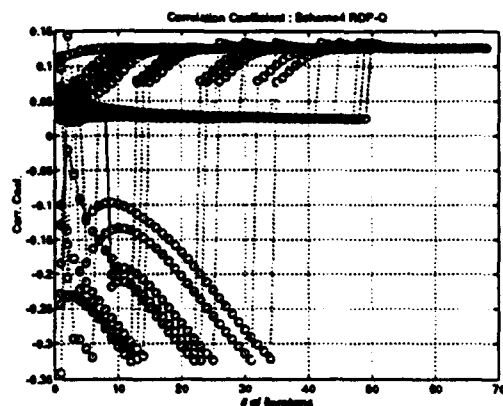
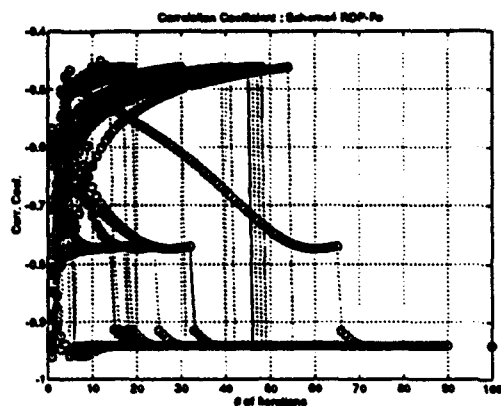


Figure 3.11 Correlation coefficient in Scheme 4.
 (a) Ψ_{∞} and corner frequency, (b) Ψ_{∞} and Q , and (c) corner frequency and Q .

Table 3.1 A posteriori Analysis

Scheme	Number of test	Correlation Coefficient			Condition Number	Estimates		
		RDP-Fo	RDP-Q	Fo-Q		RDP	Fo	Q
1	30	-0.83	0.36	-0.76	3250	1.00	1.50	50.0
2	30	-0.96	0.22	-0.46	8530	1.00	1.50	50.0
3 (D=10)	30	-0.89	0.13	-0.43	242	1.00	1.50	50.0
4 (D=10)	30	-0.94	0.12	-0.40	282	1.00	1.50	50.0
4 (D=300)	5	-0.59	-0.03	-0.38	29	1.00	1.50	50.0

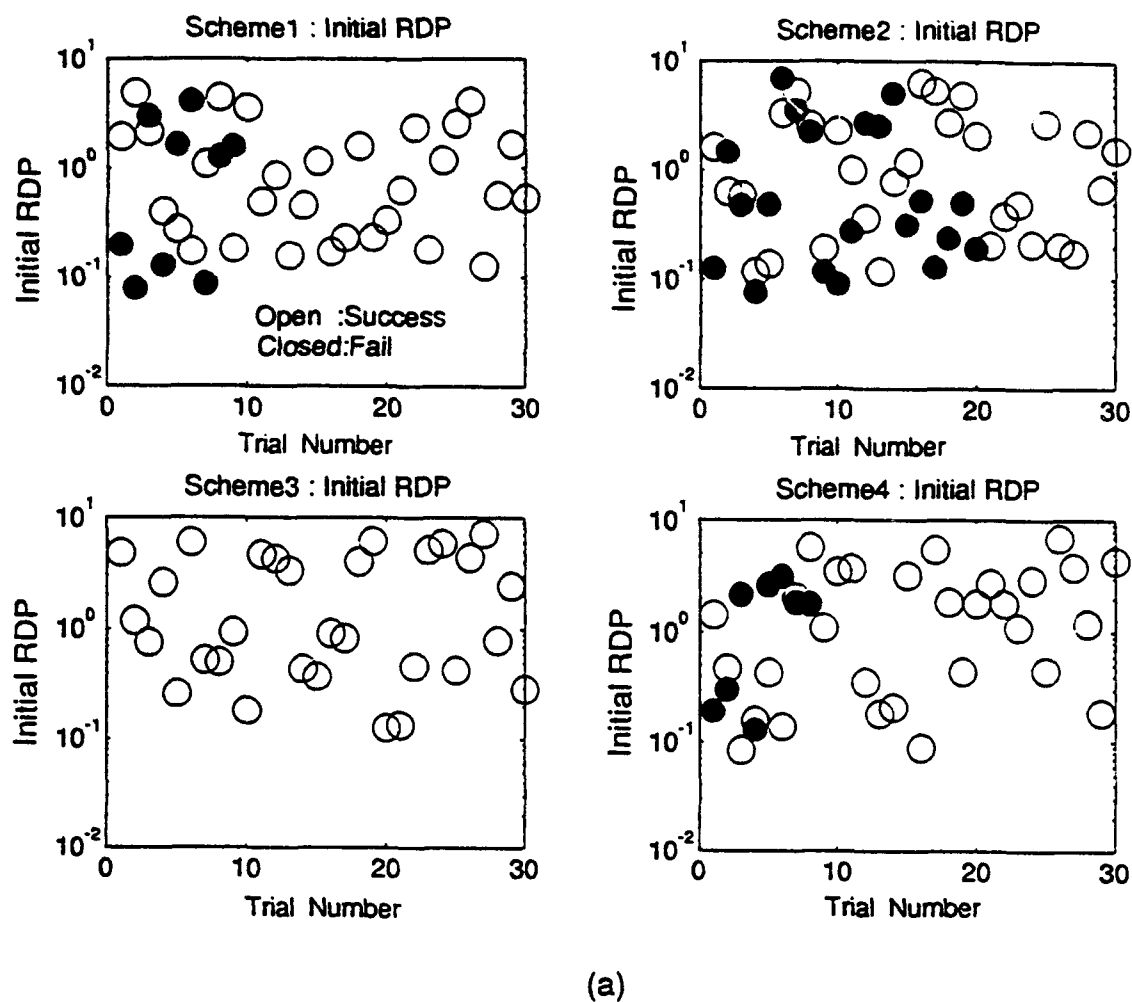
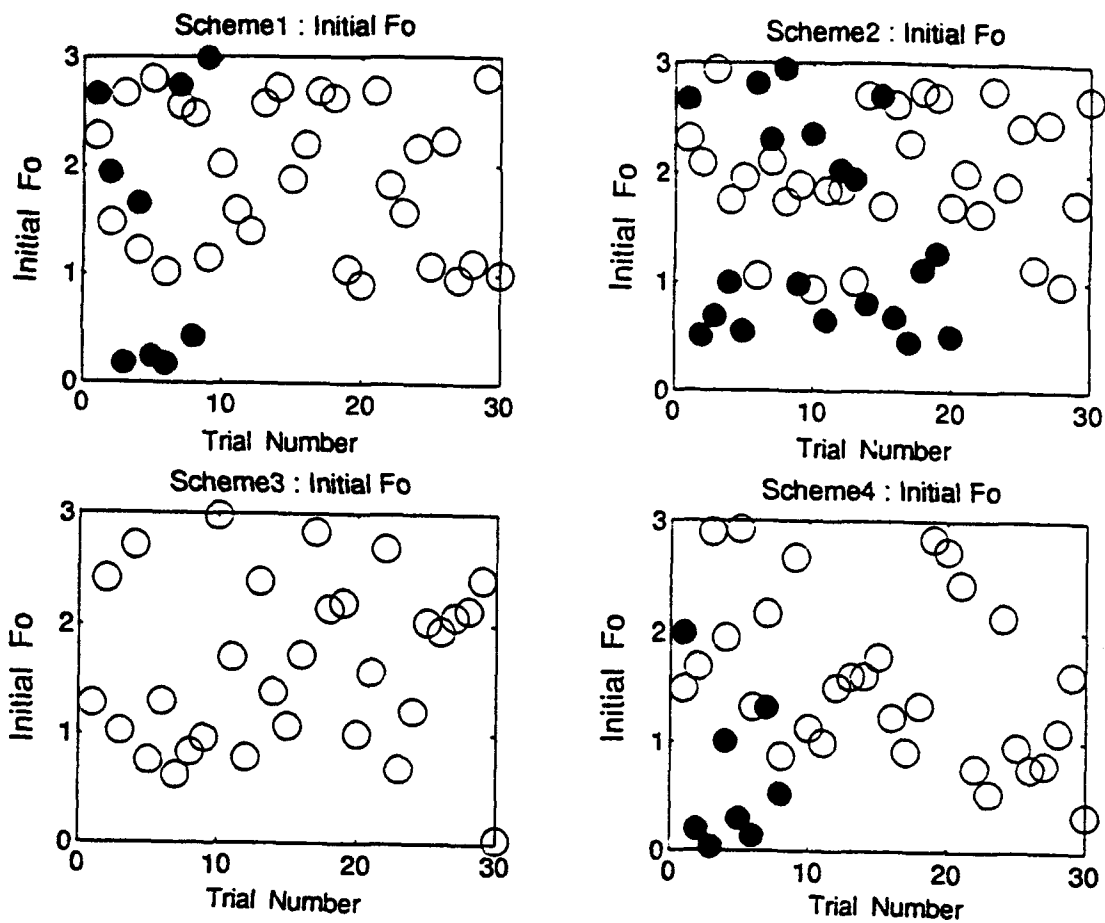


Figure 3.12 Rate of failure. There is no clear distinctive criteria for a success or a failure by the initial values of (a) Ψ_{∞} and (b) corner frequency. Small Q , however, causes divergence of inversion in (c).



(b)

Figure 3.12 Continued.

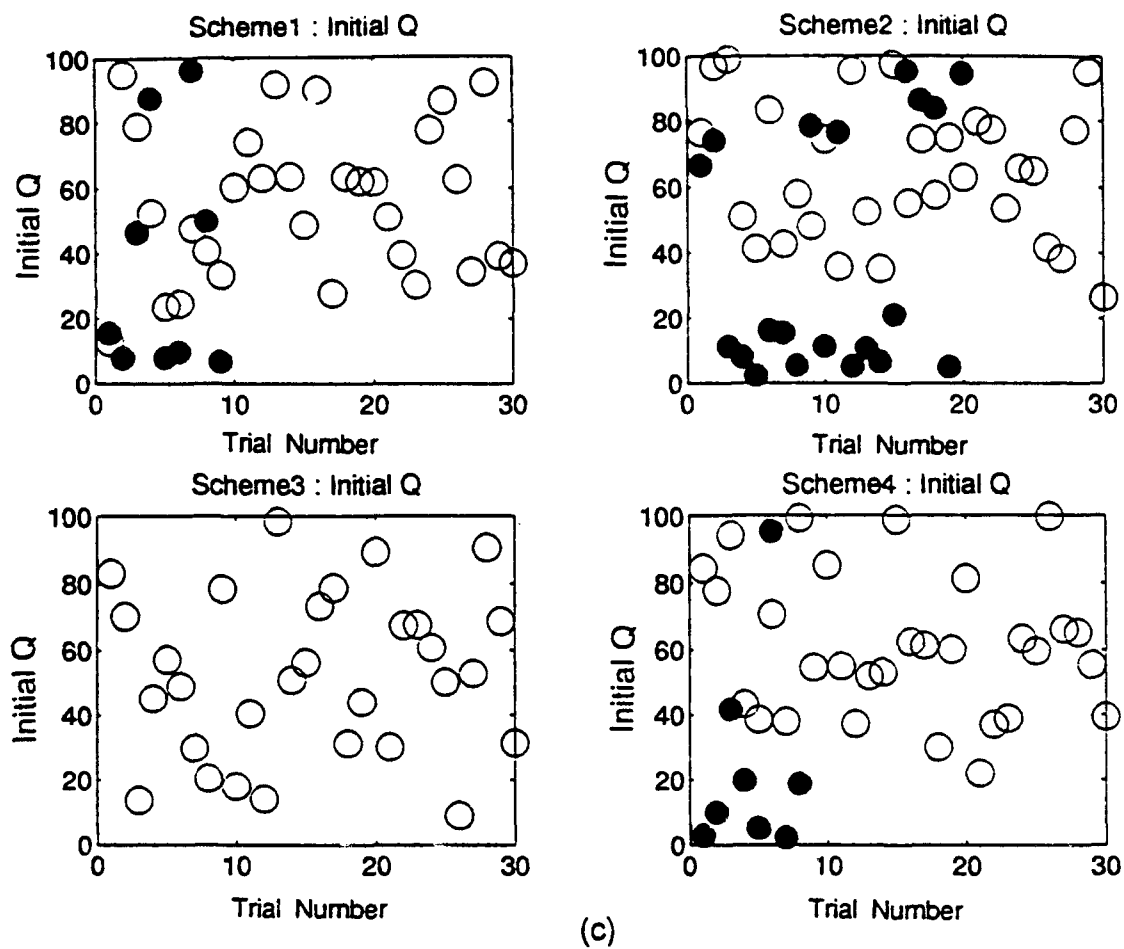


Figure 3.12 Continued.

which the model starts to converge to the expected value while the closed circles denote the failure by divergence. The distribution of the initial values illustrate that the divergence can occur when small Q (Figure 3.12, c) is given as an initial value for the prewhitening schemes (Scheme 2 and 4) while there are no clear pattern in other parameters (Figure 3.12, a and b). It is, therefore, safe to input slightly larger value for Q than expected. It is also shown that Scheme 3 (damping only) is always convergent from any arbitrary initial value. It can be understood by the trade-off between the rate of convergence and the rate of failure since most of the failure is occurred by the divergence of the parameter. Large changes in parameters between inversions increase the possibility of divergence. If it converges, however, it is fast. The rate of change of parameters is determined by the scheme itself, relaxation parameter and the applied damping parameter.

Figure 3.13 shows the relation between a given initial input and the total number of iterations. Scheme 4 with small damping parameter (10) was used in this test. There is a general trend for the number of iterations to increase as the deviation from the expected values increases for Ψ_{∞} and f_0 plot. Initial Q value, however, does not significantly affect the rate of convergence. Absolute deviation in these examples was taken as the difference of the initial value and

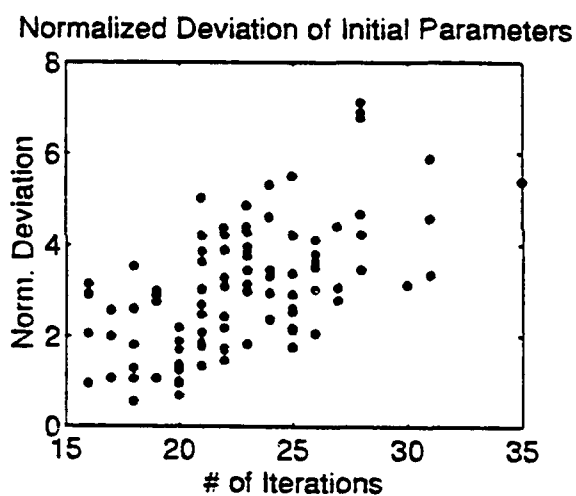
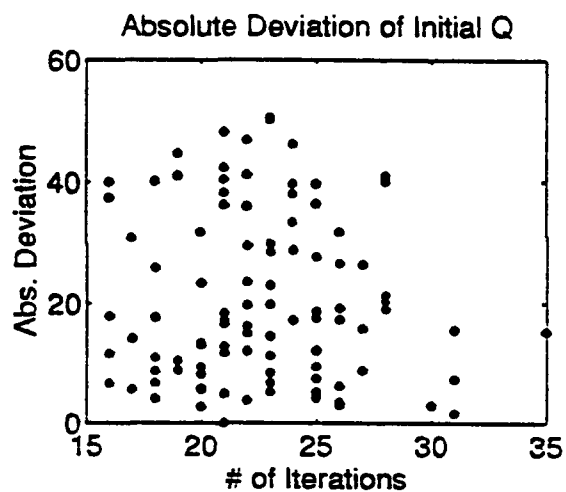
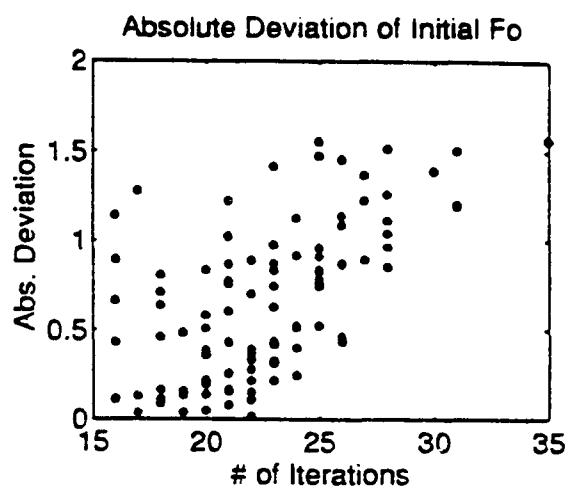
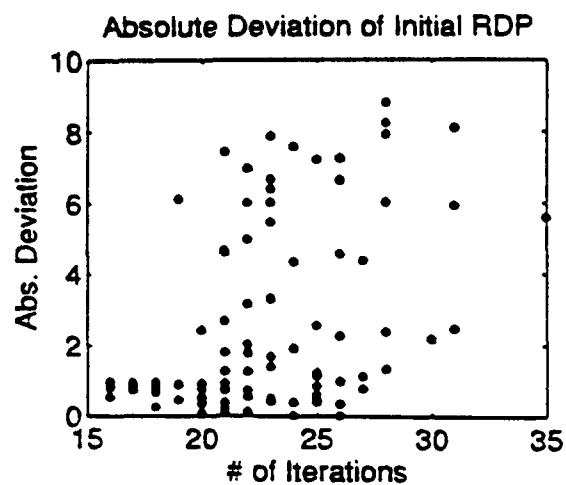


Figure 3.13 The closer the initial guess to the expected values, the faster the convergence.

the estimated parameter. Normalized deviation, in the last plot, is calculated as

$$N.D. = \sum_{i=1}^3 \frac{(I_i - E_i)}{E_i} \quad (i=1,2,3)$$

where I_i : initial value of parameter i
 E_i : estimated value of parameter i

Normalized deviation shows a clearer dependence between the initial guess and the total number of iterations necessary to reach the termination criteria.

A similar test was performed for data with theoretical random noise. The inversion shows biased steady state RDP estimation with all schemes (Figure 3.14, a and also Table 3.2). It is not surprising that steady state RDP shows biased result because noise in a few low frequency values is sufficiently large to affect the estimate. This type of bias of steady state RDP is common in the actual data and is most troublesome in estimating true steady state RDP. The method to estimate more reliable parameters from the limited data including noise will be discussed and tested empirically later in the method with bootstrap. Strong bias of Q in Scheme 1 and 3 (both, without prewhitening) is mainly due to the decreased weight of the high frequencies in these schemes (Figure 3.14, c). Noise at low- and mid-frequencies are responsible in this case. It is, therefore, imperative

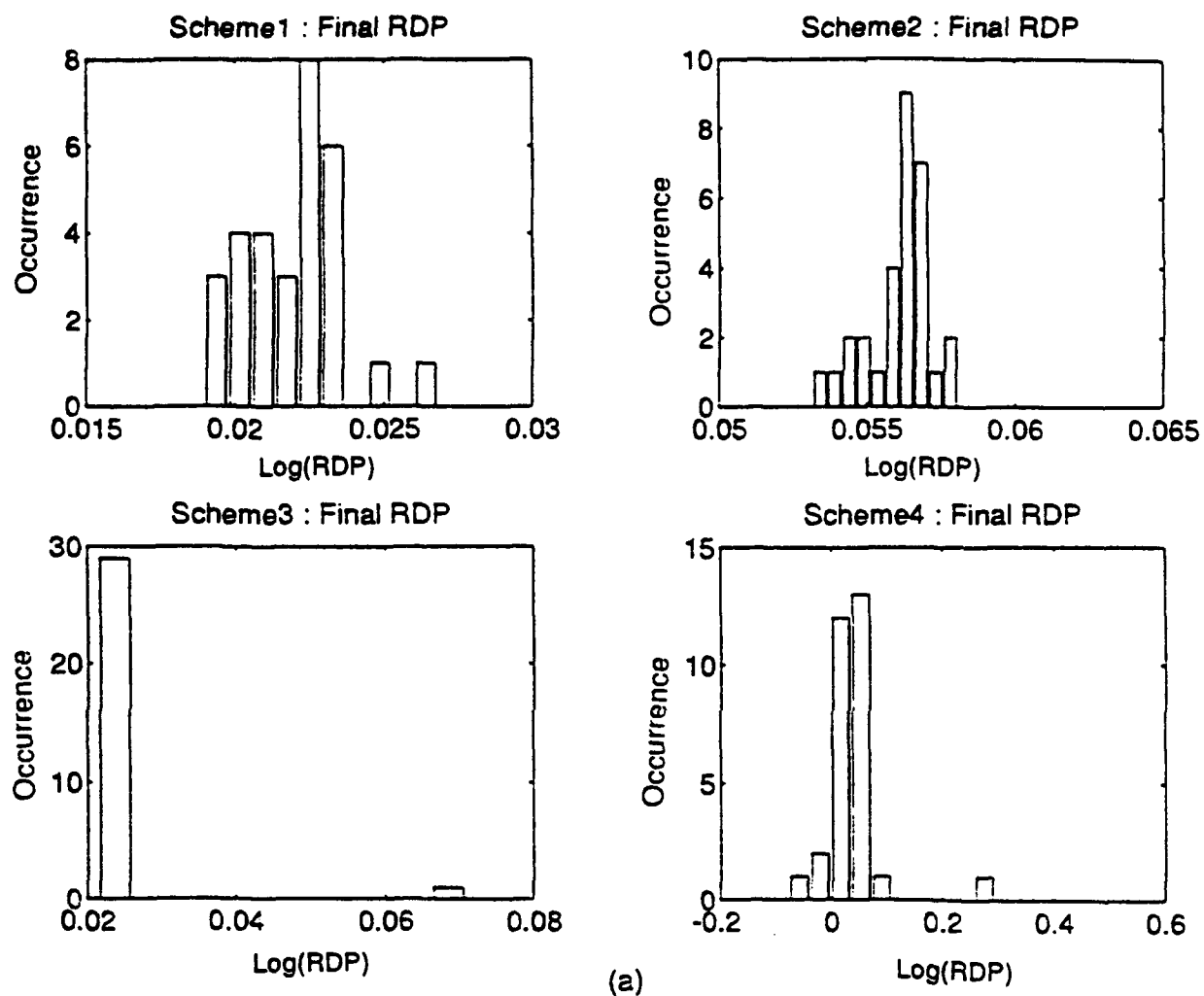


Figure 3.14 Estimation of parameters from various schemes. Ideal data including theoretical noise are used. Parameters are obtained from 30 different initial values. (a) Ψ_{∞} ; (b) corner frequency; (c) Q .

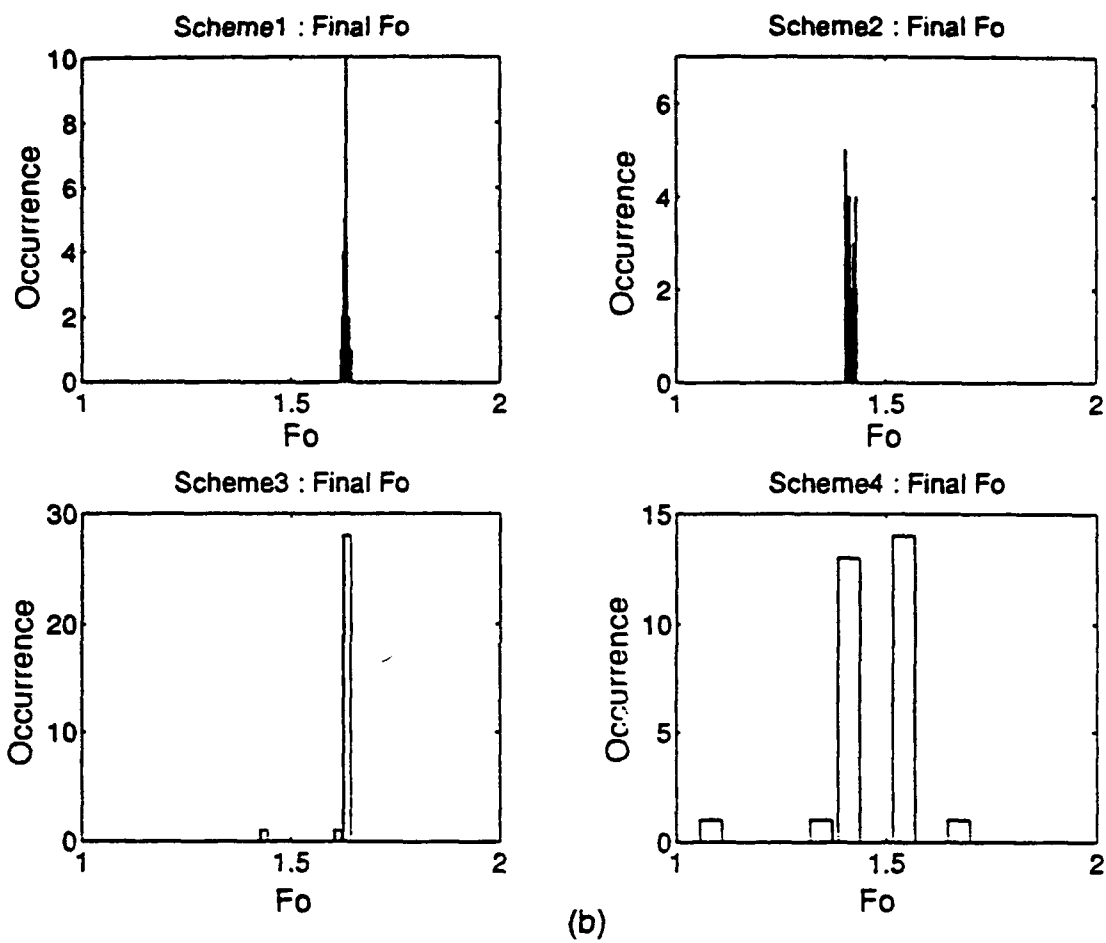
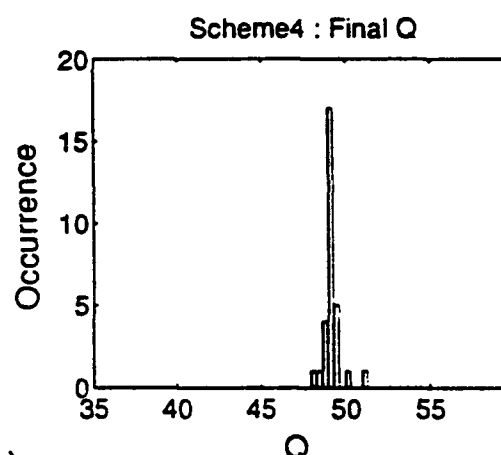
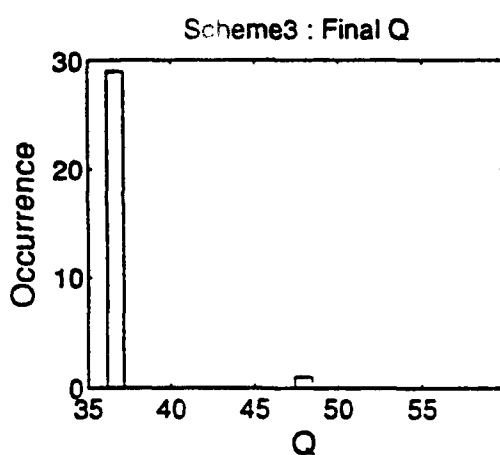
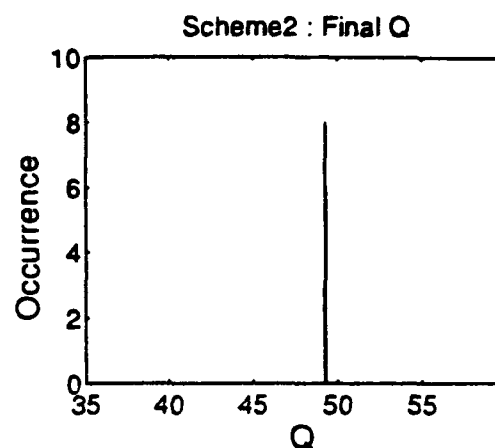
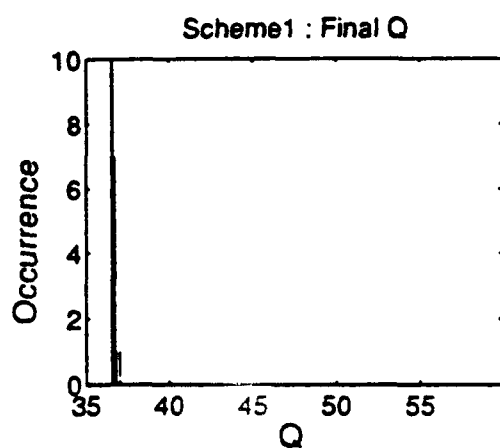


Figure 3.14 Continued.



(c)

Figure 3.14 Continued.

to prewhiten the data to resolve the attenuation effect when noise is included in the data. All schemes except 4 lead to biased corner frequency estimates (Figure 3.14, b). The bias in corner frequency (f_0) seems result from trade-off between Ψ_∞ and corner frequency and between corner frequency and Q . These trade-offs will be discussed fully later in the posterior analysis of correlation coefficient derived from the covariance matrix.

Figure 3.15 displays the number of iterations in each scheme. The histograms show the same pattern as those with ideal data. Compared to the ideal case, the median value of iteration in each scheme is increased with an addition of noise. The initial value plots do not show any peculiar difference from the ideal data (Figure 3.16). However, the ratio of failure is higher than with ideal data. Divergence generally occurs when small a Q was given as an initial value. Scheme 3 is still stable for all initial values. Scheme 3 has the smallest step length and is most likely to converge.

Figure 3.17 displays a set of resolution matrices through a whole set of iterations. From the top-left to the right, to the bottom-right, a resolution matrix is changed as the iteration proceeds. Each box in the figure represents 3×3 matrix, whose row and column corresponds to steady state RDP, corner frequency and Q respectively. The value of each element is converted to 64 levels of gray scale (black=1,

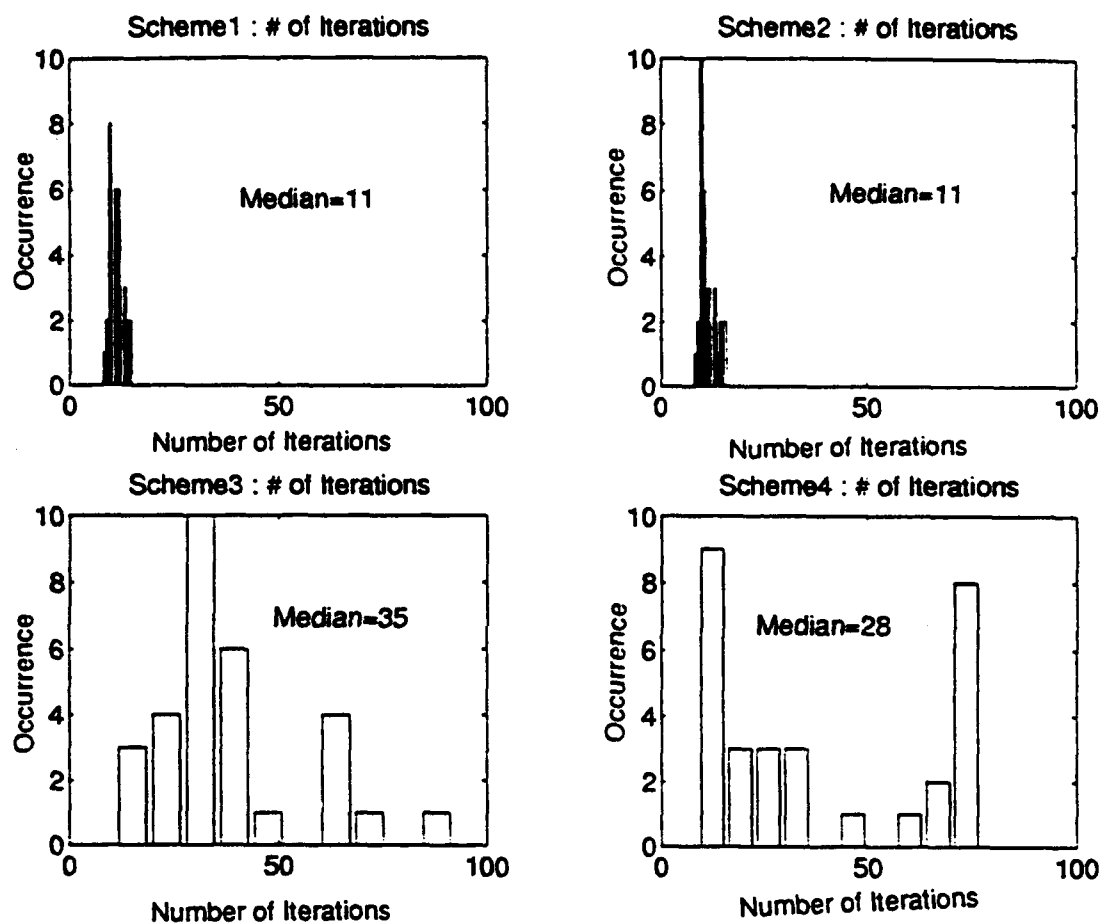
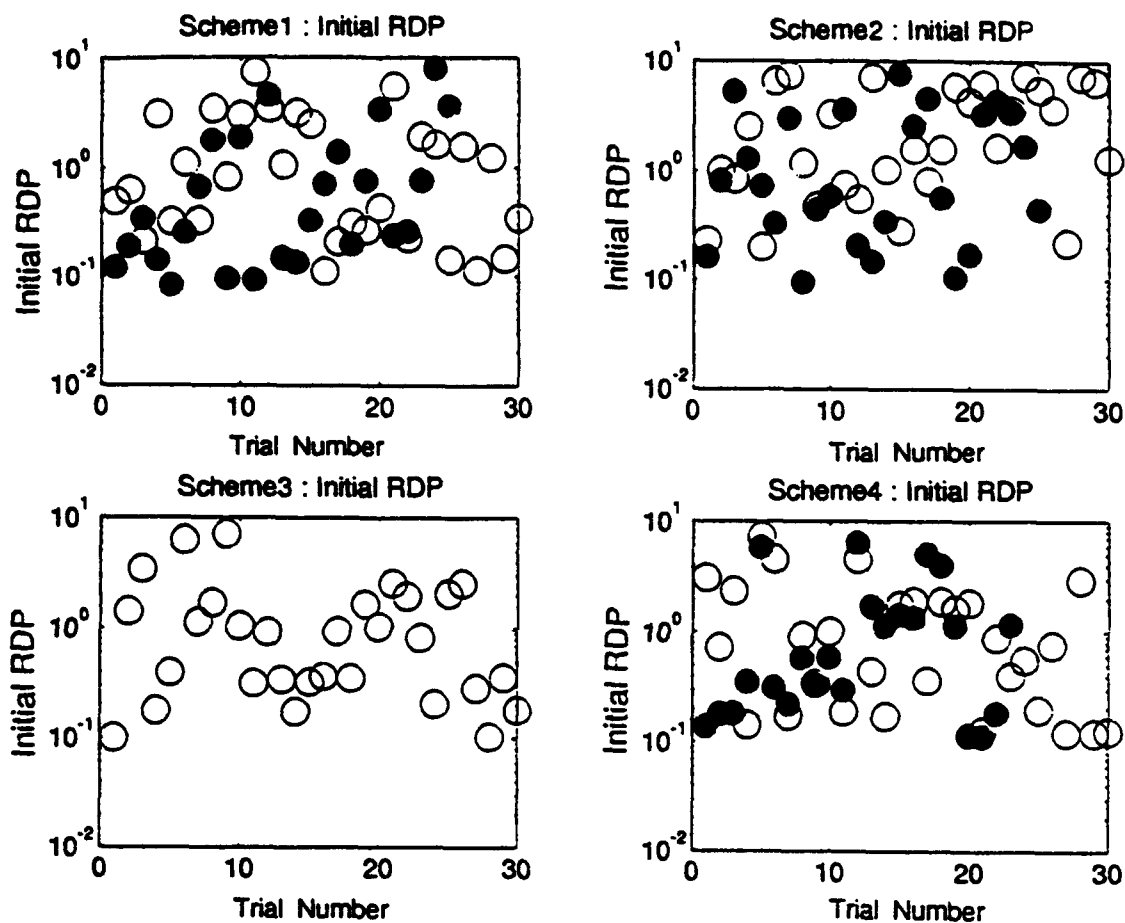
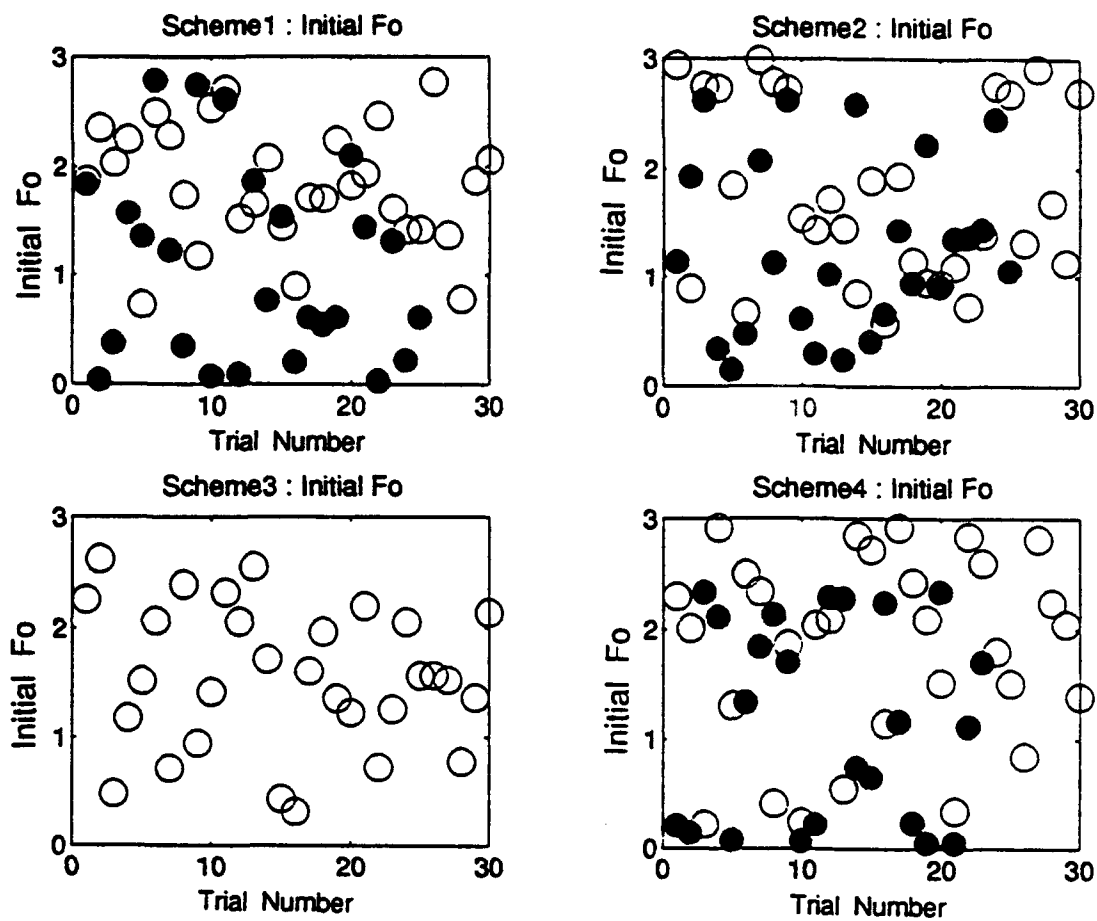


Figure 3.15 Number of iterations taken from each scheme. Number of iterations are increased by the noise effect.



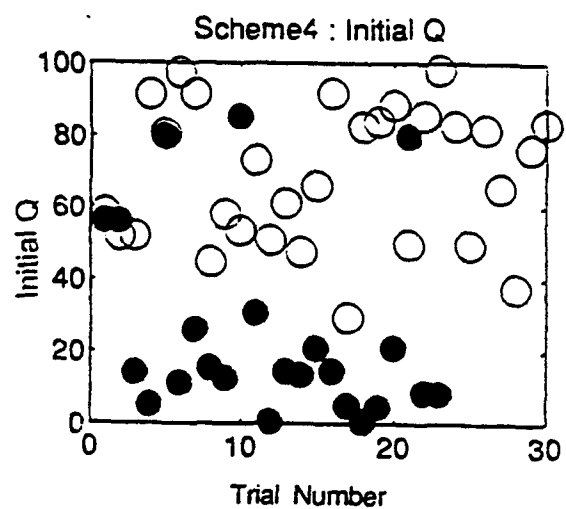
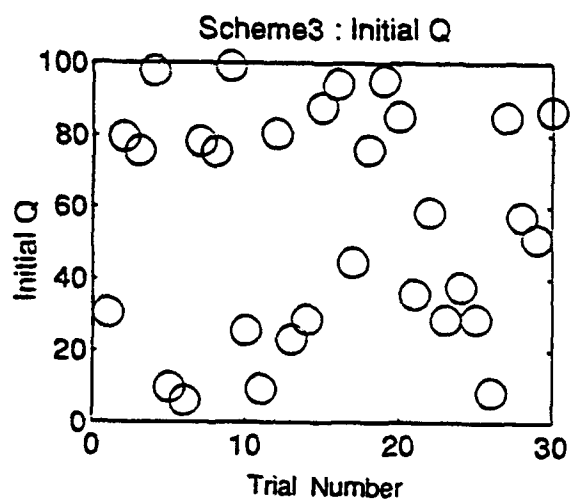
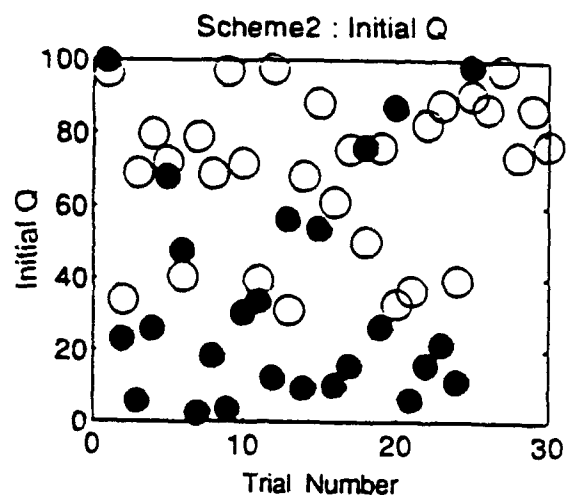
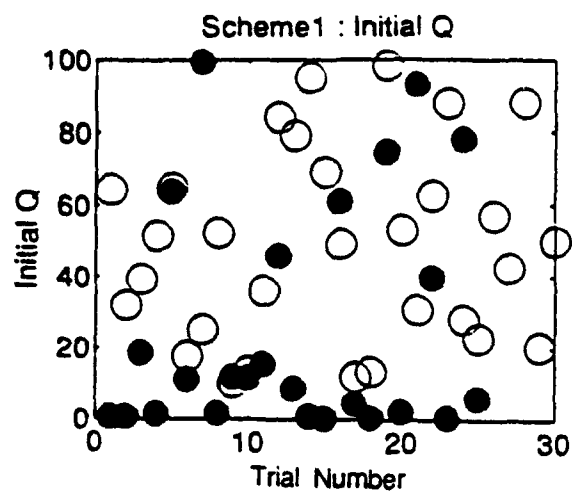
(a)

Figure 3.16 Rate of failure. Divergence generally occurs when the initial value of Q is too small. The rate of failure is raised by the noise effect.
 (a) Initial steady state RDP; (b) Initial corner frequency; (c) Initial Q



(b)

Figure 3.16 Continued.



(c)

Figure 3.16 Continued.

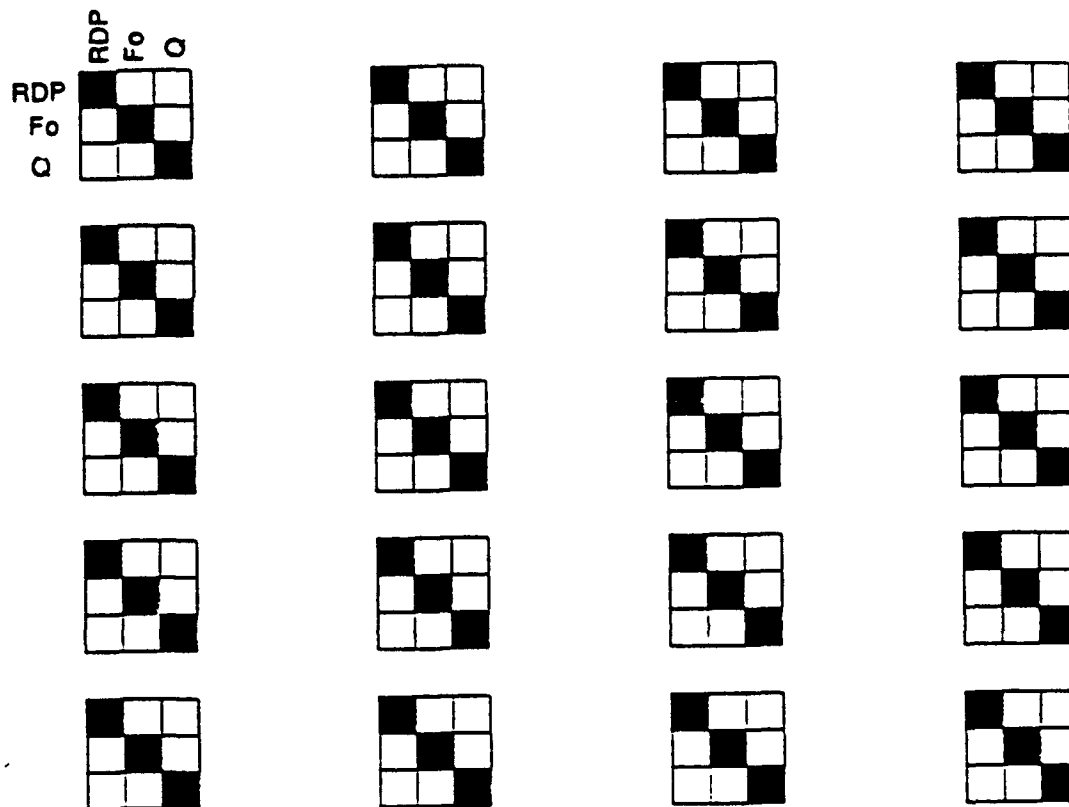


Figure 3.17 Resolution matrices. Each row and column corresponds to the resolution between parameters. The resolution matrices are calculated at each iteration from the top-left, to the bottom-right. Resolved kernels are shaded in 64 levels of gray (0=white, 1=black). (a) Scheme 1; (b) Scheme 2; (c) Scheme 3; (d) Scheme 4.

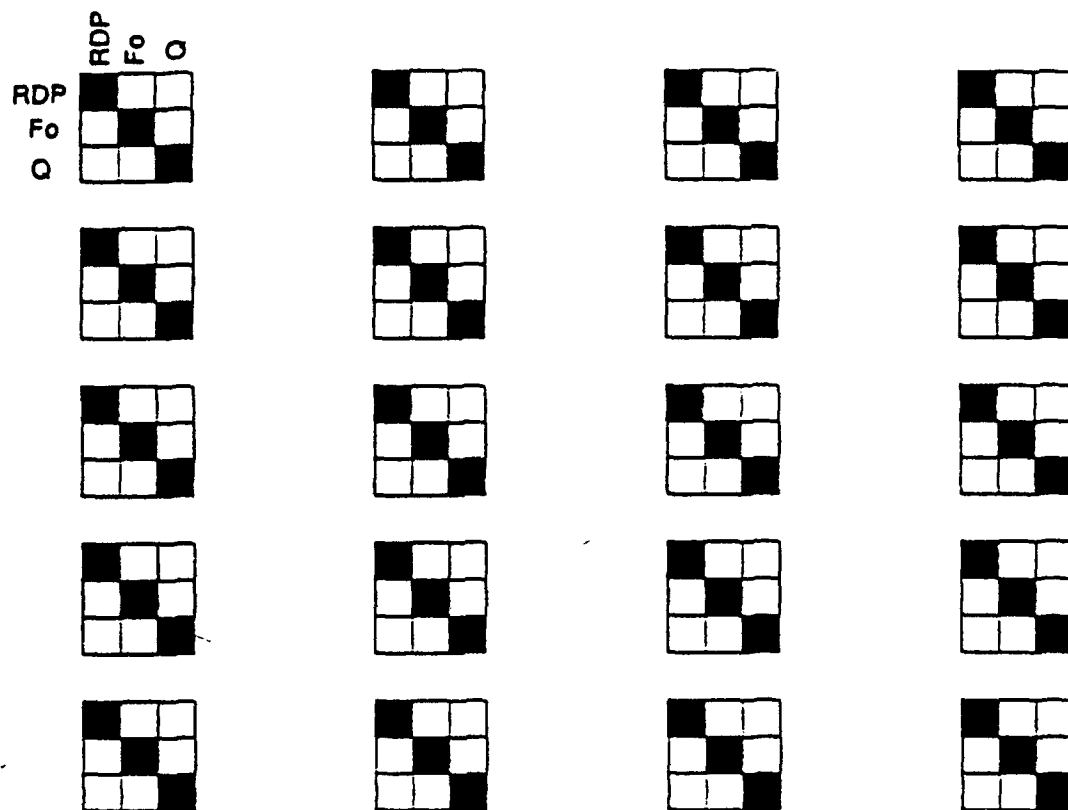


Figure 3.17 Continued.

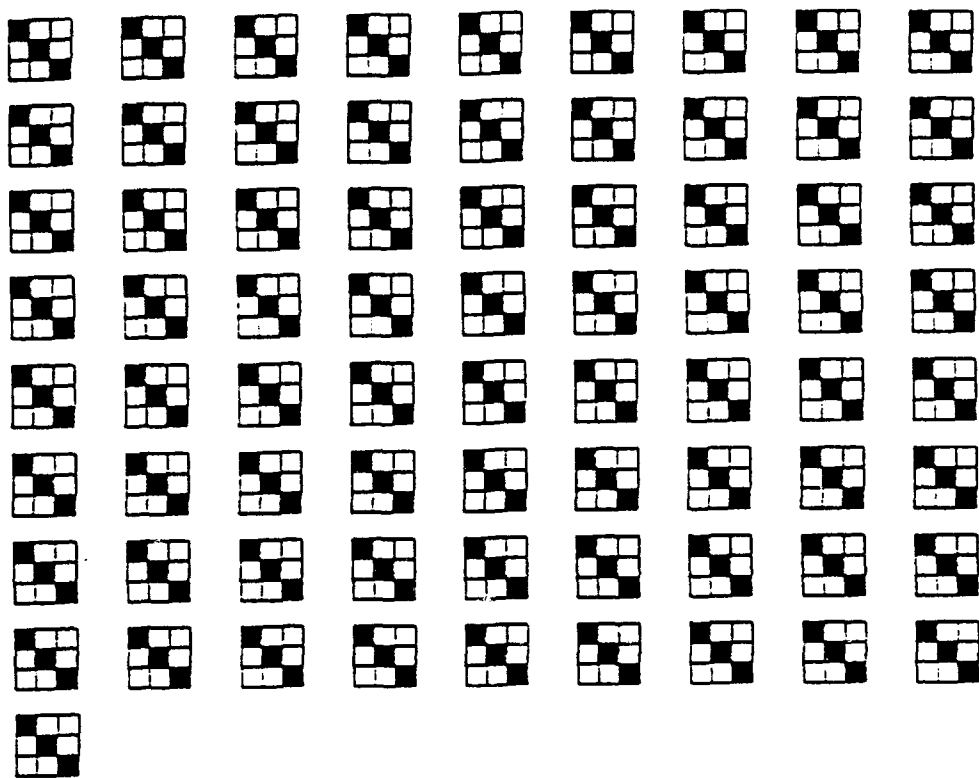


Figure 3.17 Continued.

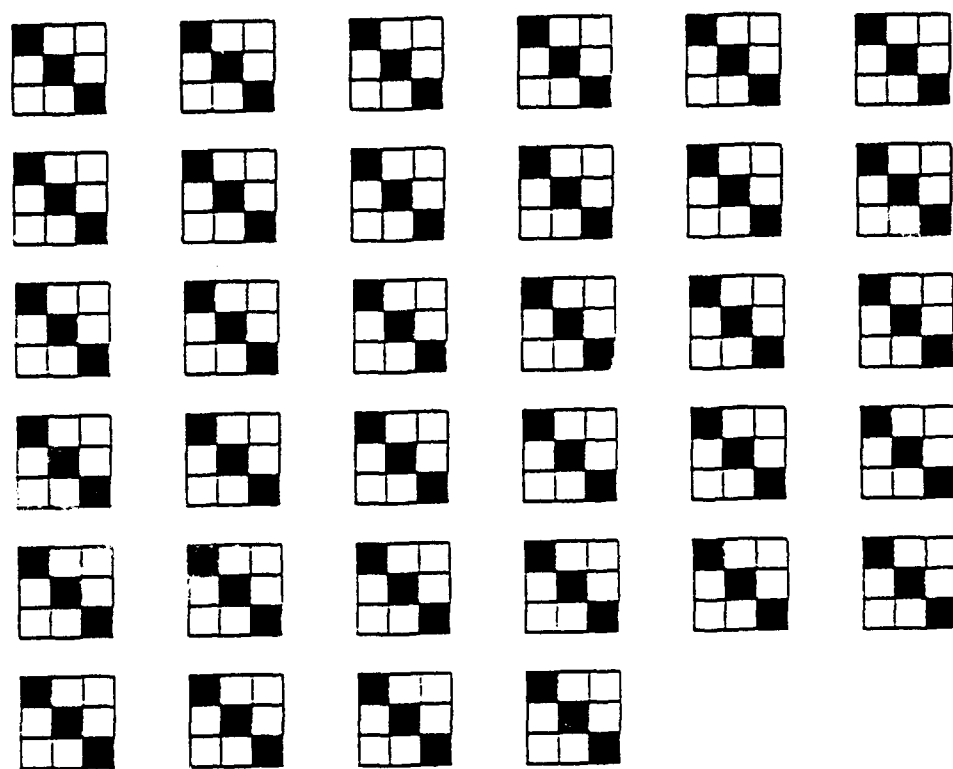


Figure 3.17 Continued.

white=0). Posterior resolution matrices(Figure 3.17, a, b, c and d) are nearly identical to the identity matrices which guarantee the uniqueness of the solution(diagonally black) for all schemes through the whole process of iteration.

Correlation coefficients show a high linear relationship between steady state RDP and corner frequency(Figure 3.18 and Table 3.2) in all schemes. It implies that two parameters trade-off with each other. The high correlation between steady state RDP and corner frequency was expected from the high frequency approximation of the forward model(Equation 2-20). Correlation coefficients between steady state RDP and Q , on the other hand, show strong independence in all cases. The degree of dependency of corner frequency and Q is between those of $\Psi_{\infty}-Q$ and $\Psi_{\infty}-f_0$. Correlation coefficients between steady state RDP and corner frequency in Scheme 3 and 4 are almost the same as the schemes without damping. This linear dependency can be reduced by introducing a larger damping parameter. Figure 3.19 shows the correlation coefficient and condition number when Scheme 4 with large damping parameter(300) was used. The correlation coefficient and condition number are stabilized with the severe penalty of drastically increased iteration. The estimates show some biases due to the noise especially at low frequencies.

If the ideal data are contaminated by the background noise(Table 3.3), the estimated parameters are different from the expected values in all schemes. Note that these biases

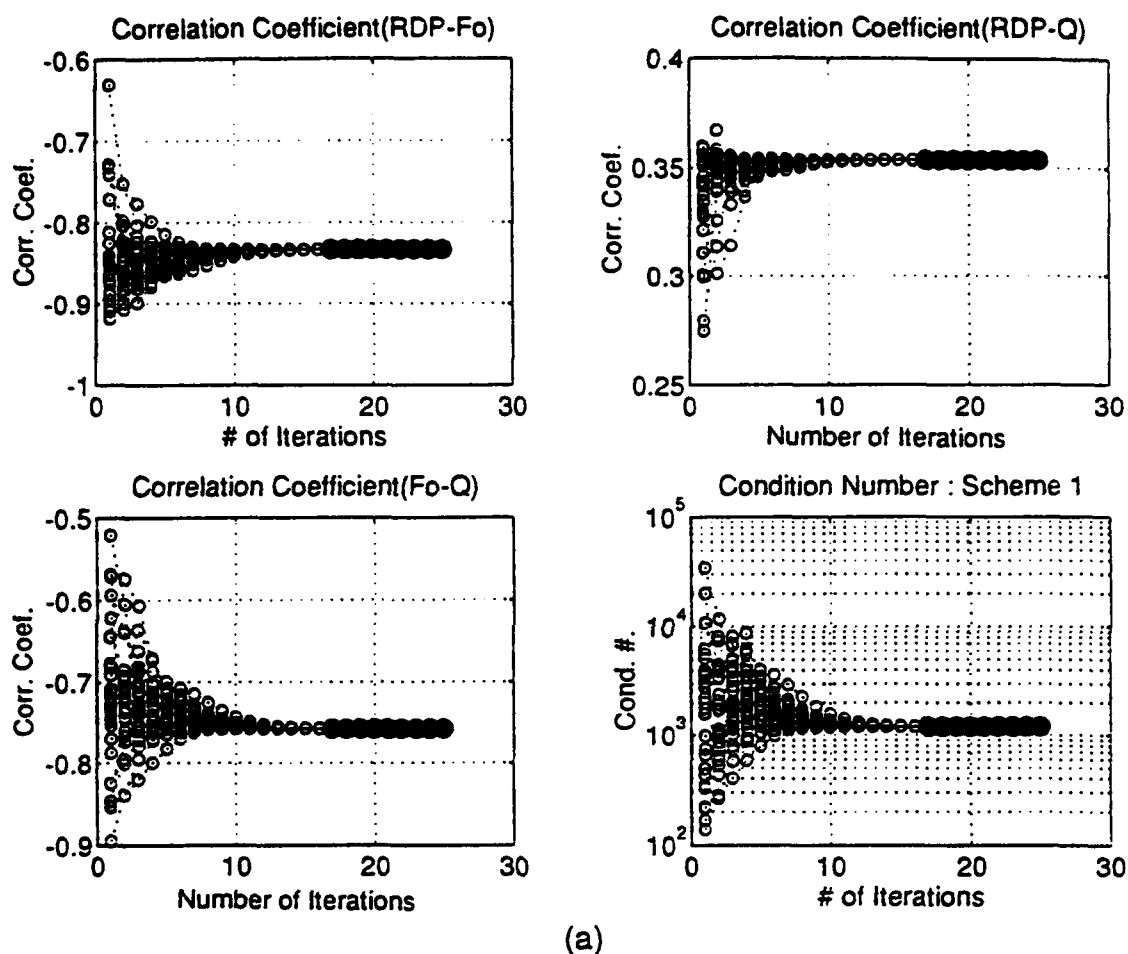


Figure 3.18 Correlation coefficients and condition numbers at each scheme with data contaminated by random noise. (a)Scheme 1; (b)Scheme 2; (c)Scheme 3; (d)Scheme 4(damping=10).

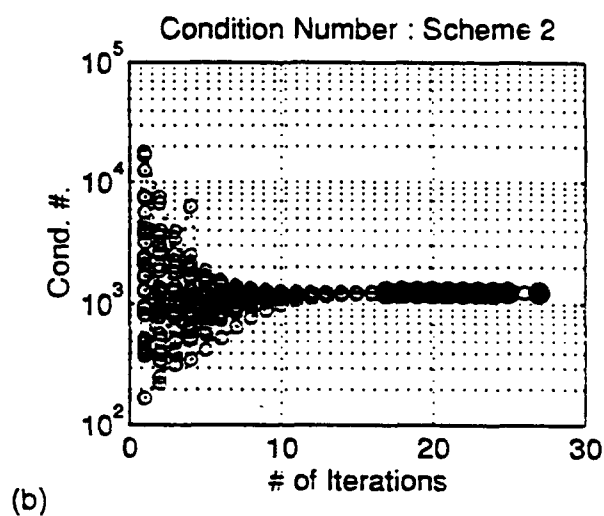
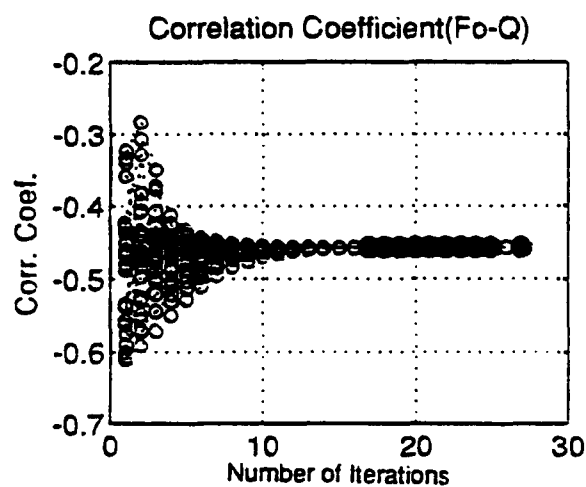
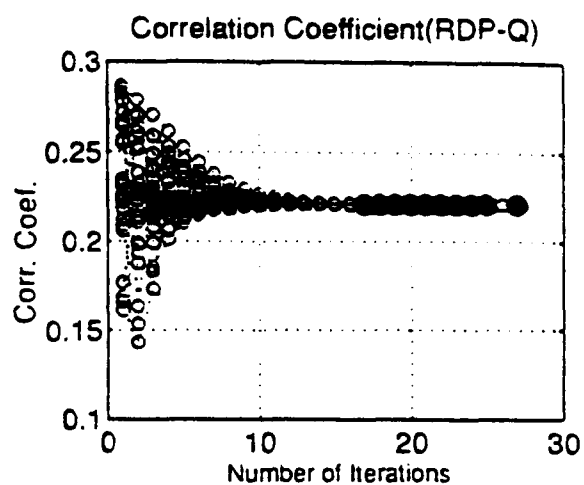
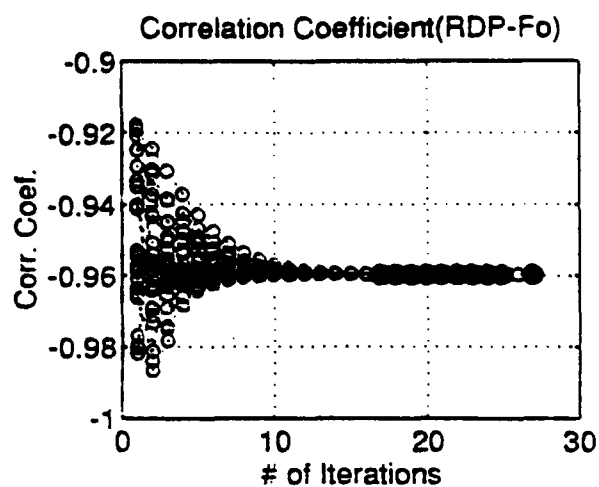
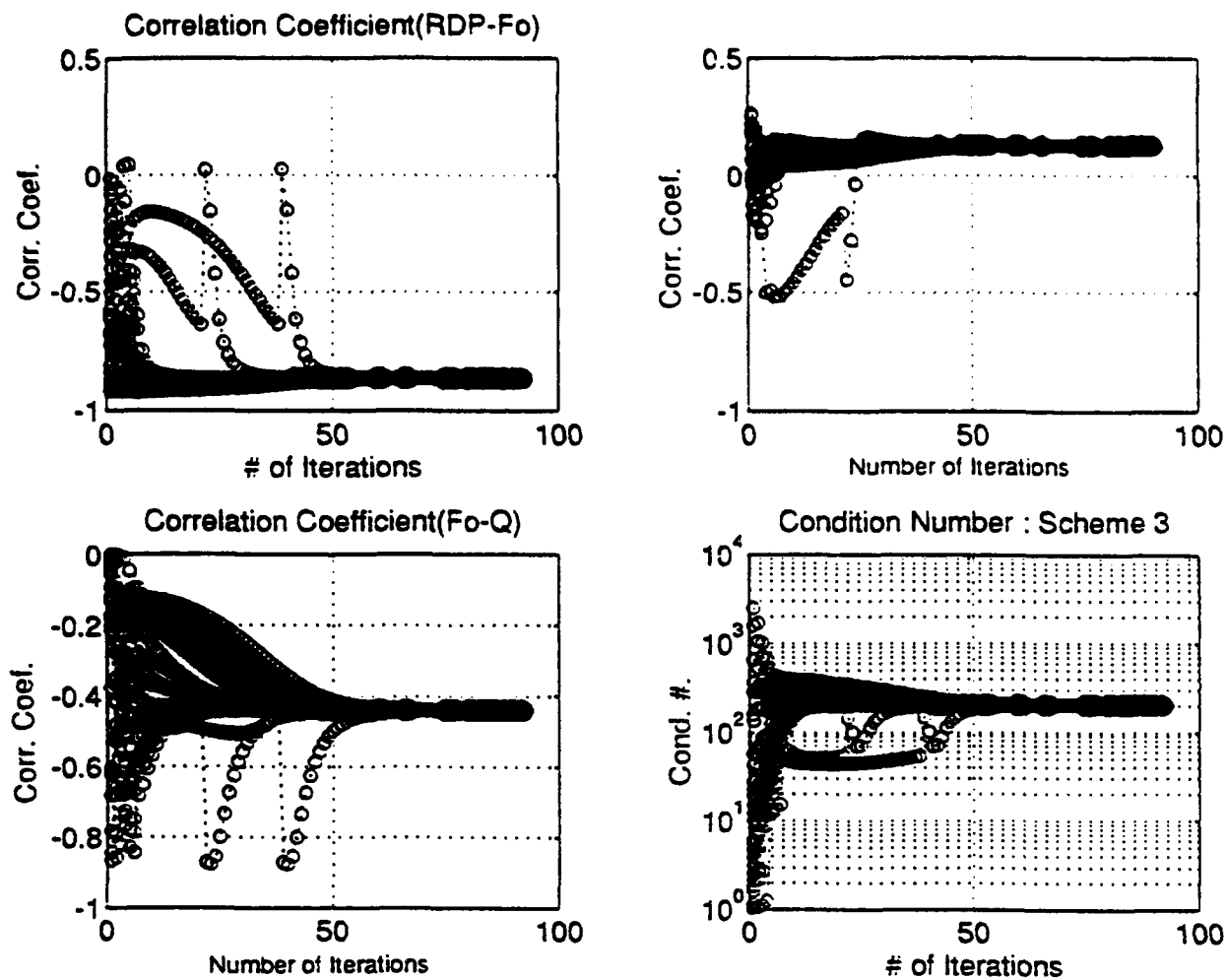
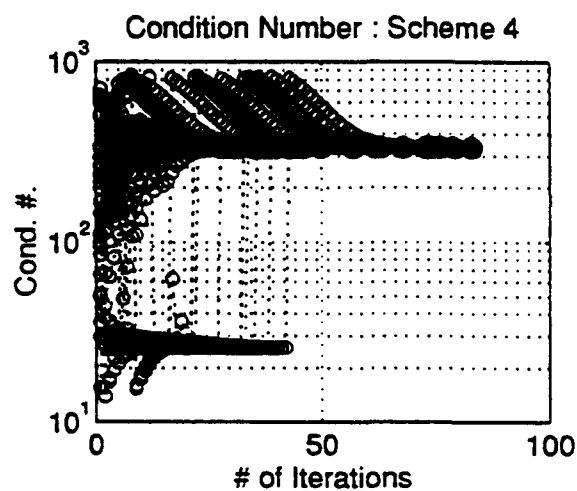
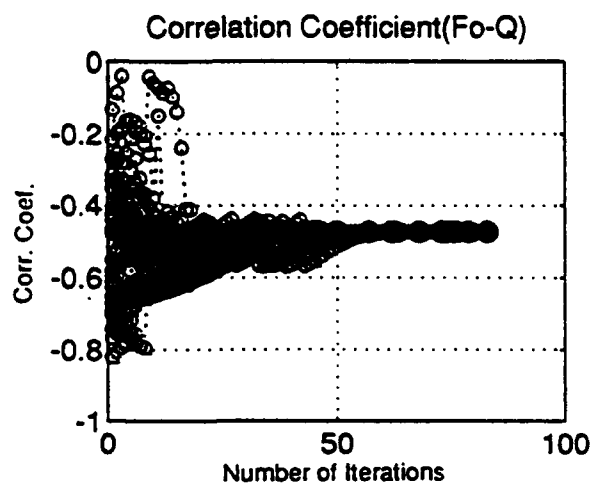
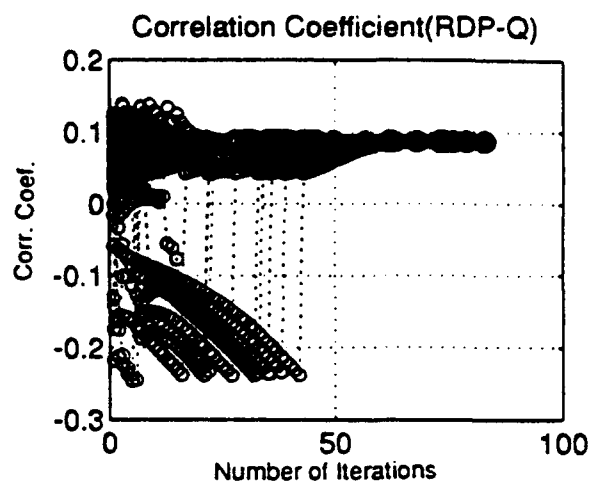
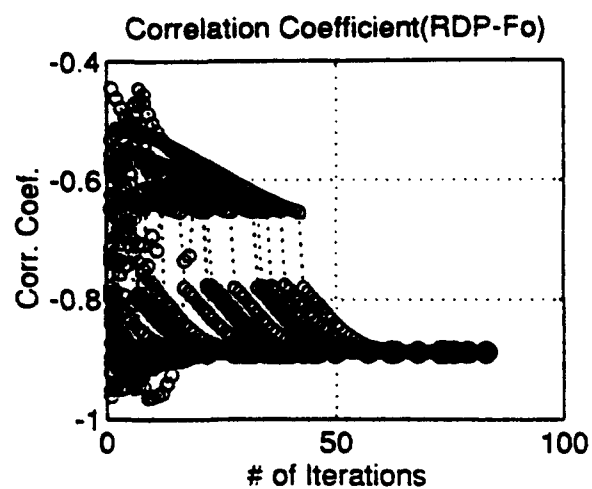


Figure 3.18 Continued.



(c)

Figure 3.18 Continued.



(d)

Figure 3.18 Continued.

table3.2

Scheme	Number of test	Correlation Coefficient			Condition Number	Estimates		
		RDP-Fo	RDP-Q	Fo-Q		RDP	Fo	Q
1	30	-0.83	0.36	-0.76	1200	1.05	1.63	36.7
2	30	-0.96	0.22	-0.46	1247	1.14	1.42	49.3
3 (D=10)	30	-0.88	0.14	-0.42	206	1.06	1.63	37.2
4 (D=10)	30	-0.89	0.09	-0.47	337	1.09	1.46	49.1
4 (D=300)	5	-0.54	-0.02	-0.41	16	1.13	1.43	49.2

* Correlation coefficient and condition number were taken from the last iteration.

* Estimates were taken from the arithmetic mean.

* D denotes damping parameter.

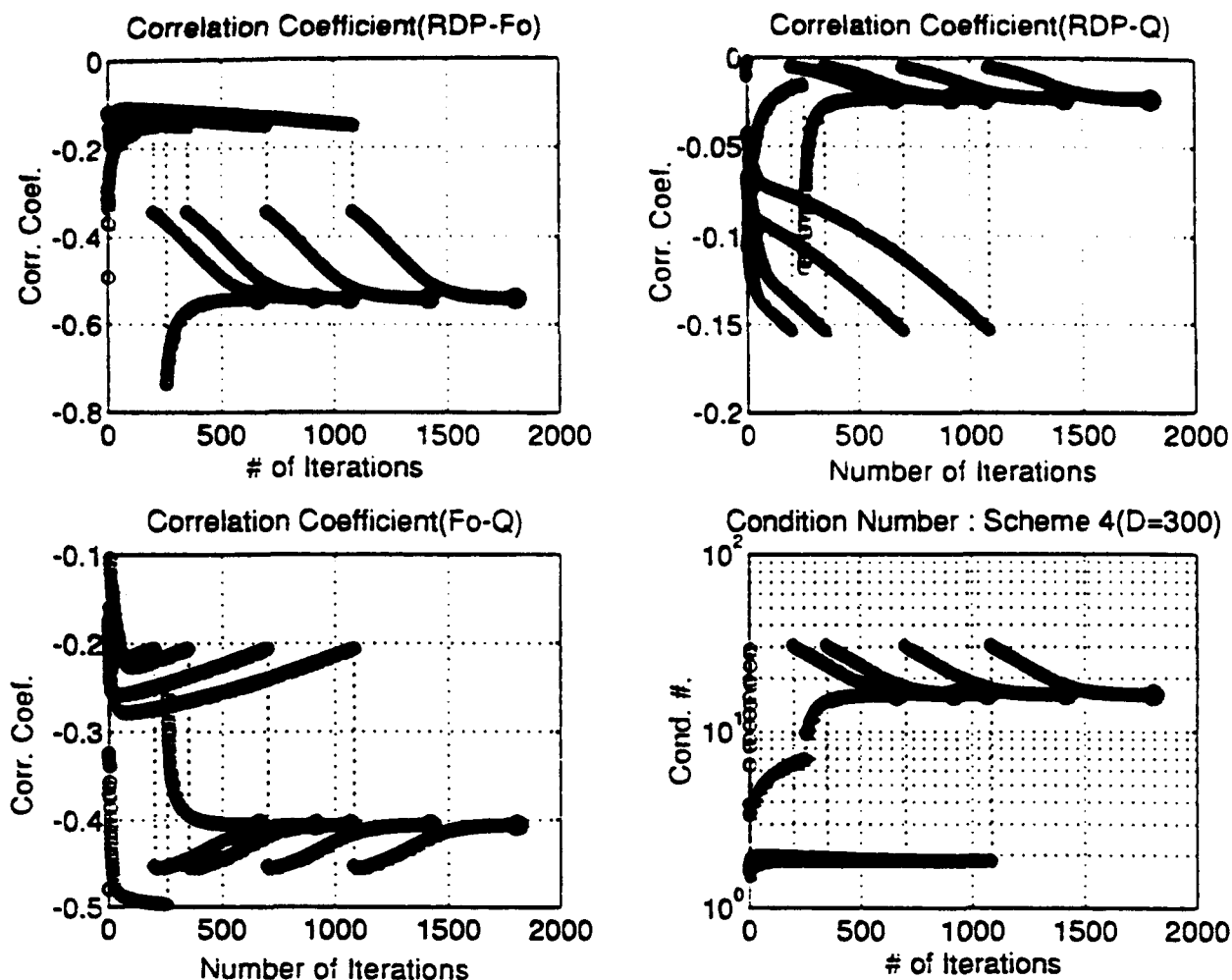


Figure 3.19 Correlation coefficients and condition numbers at Scheme 4 with large damping parameter(300). The data used for Figure 3.18 are used. Correlation coefficient between Ψ_∞ and f_0 is reduced.

table3.3

Scheme	Number of test	Correlation Coefficient			Condition Number	Estimates		
		RDP-Fo	RDP-Q	Fo-Q		RDP	Fo	Q
1	10	-0.88	0.33	-0.66	7483	1.23	1.29	61.1
2	10	-0.98	0.17	-0.34	5050	3.73	0.73	57.9
3 (D=10)	10	-0.81	0.04	-0.22	353	1.23	1.29	61.1
4 (D=10)	10	-0.86	-0.20	-0.23	48	3.73	1.73	57.9
4 (D=300)	6	-0.38	-0.20	-0.14	2.7	3.72	0.73	57.8

* Correlation coefficient and condition number were taken from the last iteration.

* Estimates were taken from the arithmetic mean.

* D denotes damping parameter.

resulted from the background noise rather than the inversion scheme. It is hard to say that Scheme 1 and 3 (both without prewhitening) show better results since source parameters may be biased by the background noise. One way to solve this problem is to maintain high signal-to-noise ratio by cutting off both sides of the spectrum. If the frequency band is limited so as to retain the high S/N ratio, the inversion output approaches to the expected values (Table 3.4). This cutoff of frequency bands is also necessary to maintain the consistency of the assumption that the distribution of the data is log-normal. The cut-off range is dependent on the S/N ratio. It is necessary to analyze the pre-event noise in an observational data to determine the reliable frequency ranges of the data. The correlation coefficient and resolution matrices do not show any apparent differences between the two data sets that could be used to make these estimates of band width. It is not always possible, however, to obtain sufficient bandwidth within which source parameters can be resolved reliably. In this case, the inversion may show ill-posedness and thus the estimates show large variation. Bootstrap will be useful to treat the data having insufficient bandwidth.

From the result of test (1) with ideal data, (2) with ideal data including theoretical noise, and (3) with ideal data including background noise, Scheme 1 (neither prewhitening nor damping) and Scheme 2 (prewhitening) are not

table 3.4

Scheme	Number of test	Correlation Coefficient			Condition Number	Estimates		
		RDP-Fo	RDP-Q	Fo-Q		RDP	Fo	Q
1	10	-0.88	0.34	-0.67	5826	1.17	1.35	57.9
2	10	-0.97	0.23	-0.44	12756	1.25	1.30	55.8
3 (D=10)	10	-0.82	0.05	-0.24	315	1.17	1.35	57.9
4 (D=10)	10	-0.66	0.03	-0.66	490	1.25	1.30	55.8
4 (D=300)	10	-0.24	-0.01	-0.39	18	1.25	1.30	55.8

* Correlation coefficient and condition number were taken from the last iteration.

* Estimates were taken from the arithmetic mean.

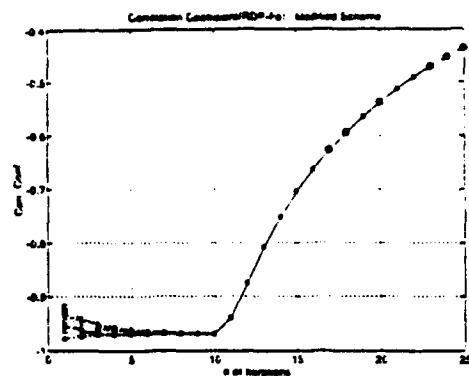
* D denotes damping parameter.

appropriate for the source parameter inversion because of the trade-off of parameter estimates despite the speed of convergence. Scheme 1 shows reliable estimates of steady state RDP and corner frequency despite of large condition number. Scheme 2 gives almost the same result as Scheme 4. Reliable estimates of source parameters despite a large condition number in both cases (Scheme 1 and Scheme 2) mean that the Brune's model is well-posed if the bandwidth of the data is sufficient. Nevertheless, Scheme 2 does not guarantee the same result if other types of source or wave propagation phenomenology such as spall, surface waves and scattering are present in the data. Scheme 3 (damping) is the worst scheme out of those tested. It is biased in its estimates as well as slow in convergence. Scheme 3 shows reliable results in estimating steady state RDP and corner frequency despite the large condition number and high linearity between steady state RDP and corner frequency, too. Scheme 4 (both prewhitening and damping) is the optimal one to resolve both source parameters and Q since this scheme is flexible, stable, and relatively fast if the damping parameterized are chosen properly. Sometimes, a set of inversions require a large damping parameter for improvement in independency between parameters. In this case, the main disadvantage of Scheme 4 is its slow speed of convergence.

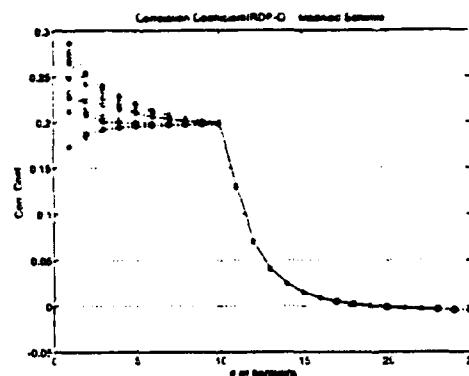
Modified scheme

It is possible to combine more than one scheme in an inverse process improve the convergence rate without sacrificing reliability. Since Scheme 2 converges fast and Scheme 4 maintains independence among parameters by an appropriate damping parameter, the new scheme adopts two processes in a single inversion. It follows the direction of Scheme 2 in a first few iterations which improves initial speed of convergence and follows the direction of Scheme 4 for maintaining independence among the parameters. The two schemes can be linked by a Hanning window so that they are connected smoothly to avoid abrupt change of damping parameter. Abrupt change of damping parameter may cause the perturbation of estimates which demerits the initial speed of convergence with Scheme 2. Actually, linking them is nothing but a Scheme 4 with varying damping parameter.

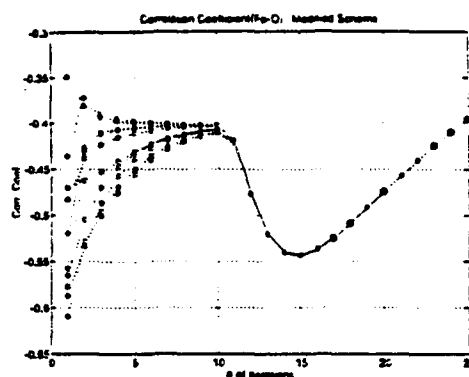
For a test of the modified scheme, ideal data with theoretical noise were used. Damping parameter was changed from 0 (no damping: Scheme 2) to 300 (Scheme 4) in ten steps. The first five iterations were performed without damping. From the sixth to fifteenth iteration, the Hanning window was applied leading to full damping (300) at the fifteenth iteration. The correlation coefficient between steady state RDP and corner frequency drops from around -0.95 to -0.55 which guarantees improvement in the linear independency between the two parameters (Figure 3.20, a). The correlation



(a)



(b)



(c)

Figure 3.20 Correlation coefficients for the modified scheme. Correlation coefficients change smoothly from the large(Scheme 2 level) to the smaller value(Scheme 4 level). (a)steady state RDP and corner frequency; (b)steady state RDP and Q; (c)corner frequency and Q.

coefficient between steady state RDP and Q drops to nearly zero (Figure 3.20, b). On the other hand, the correlation coefficient between corner frequency and Q increases from -0.45 to -0.6 (Figure 3.20, c). Even though the linear dependency is increased between corner frequency and Q , it is moderate compared with other correlation coefficients. These values of correlation coefficients are comparable to those of Scheme 4 with its large damping parameter (Table 3.2). Figure 3.21 shows the number of iterations to resolve the source parameter. The median value of number of iterations with the modified scheme, 22 iterations, is much smaller than that of Scheme 4 with large condition number (Figure 3.19), around 1000 iterations, and even smaller than that of Scheme 4 with small damping parameter (Figure 3.15), 28 iterations. Figure 3.22 illustrates how the condition number changes as the iteration proceeds. In the first few iterations, the Hessian is nearly singular (large condition number) because of the trade-offs between parameters. At this stage of Scheme 2, trade-off is mainly due to the dependency between steady state RDP and corner frequency (Figure 3.20, a). Transition from Scheme 2 to Scheme 4 by applying Hanning window results in a continuous drop of condition number to less than 100. Even though correlation coefficient and condition number did not converge sufficiently because the iteration meets the termination criteria before the modified scheme reaches the convergence region, the tests are sufficient to show that the

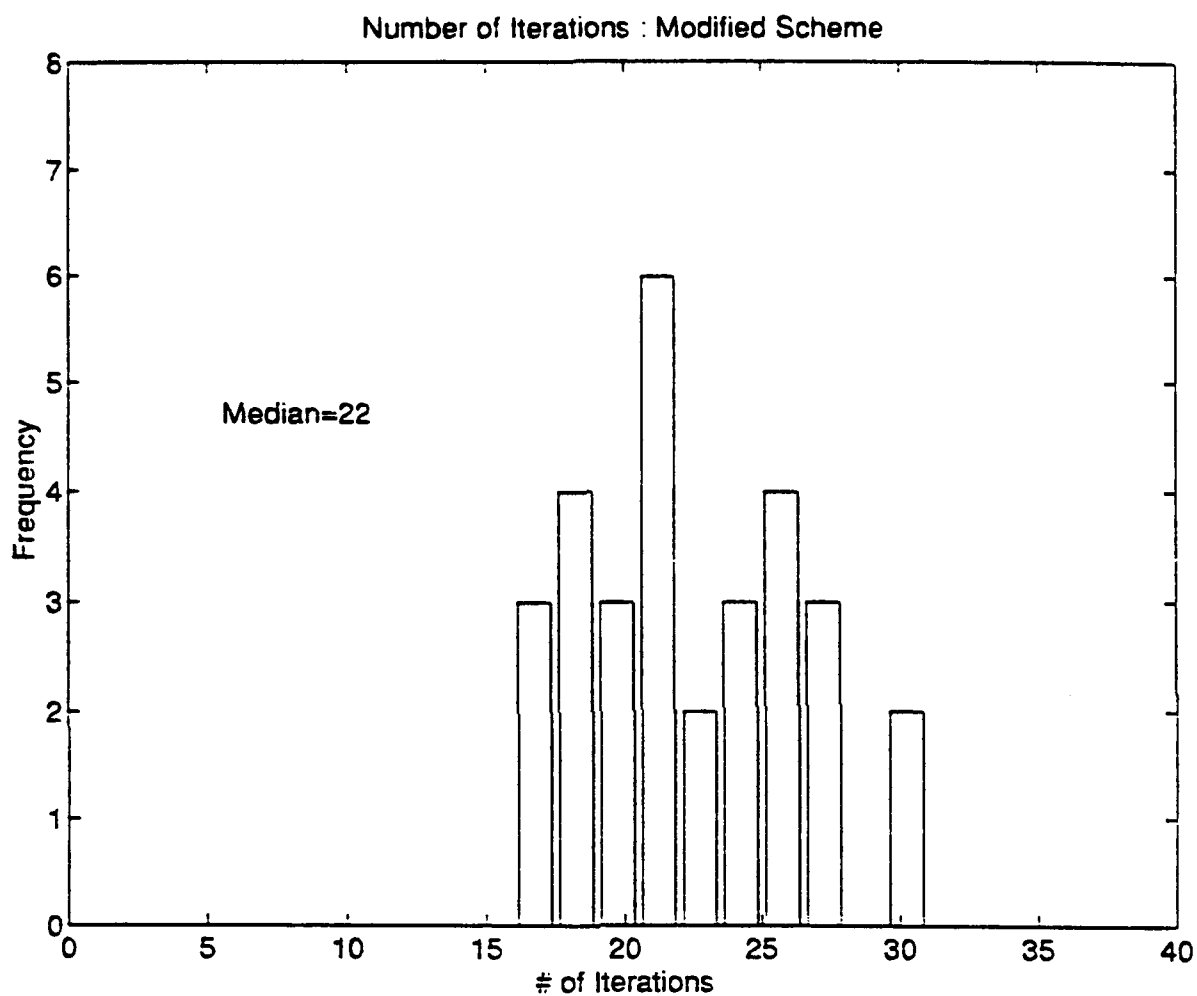


Figure 3.21 Number of iterations taken to resolve the parameters. 30 different initial values are used. The median of the number of iterations is reduced to 22 from 1000 at Scheme 4 with the same damping parameter (Figure 3.19)

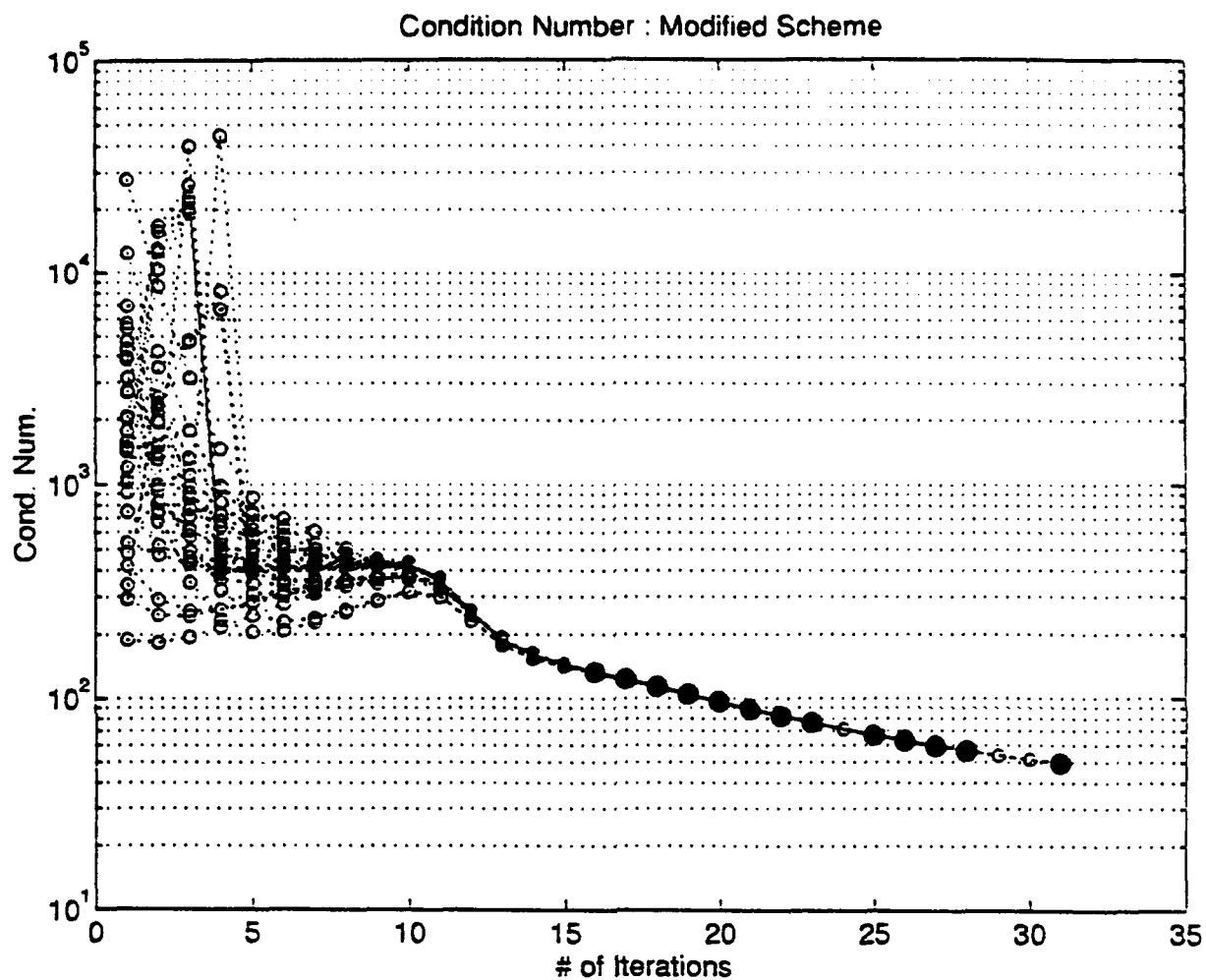


Figure 3.22 Condition numbers in the modified scheme. Condition number is dropped below 100 at the final stage of iterations.

modified scheme is faster than Scheme 4 maintaining the same degree of independency in its final step.

Prewhitening process

It is assumed that the prewhitening function(w) approaches the unknown model(g) as the iteration proceeds(Chapter 2.B.2) in order to apply the prewhitening process as a part of inverse process. Figure 3.23 shows how the residuals change as iterations proceed. Differences between the data and the estimated model converge to zero rapidly in the entire range of frequencies after a few iterations. This verifies that the initial curve which may not be appropriate as a prewhitening function approaches the final proper curve(final model) after a few iterations. Ideal data and Scheme 1 were used to verify the convergence of the prewhitening function.

Synthetic Data with Overshoot

Data for the test with overshoot were generated from the VSB model which included source overshoot. Inversion tests included (1)ideal data with theoretical noise and (2)ideal data with background noise as was done in previous section.

Inversion with the Brune's model

To investigate the possible applicability of the simple source model(Brune's model) in the resolution of source

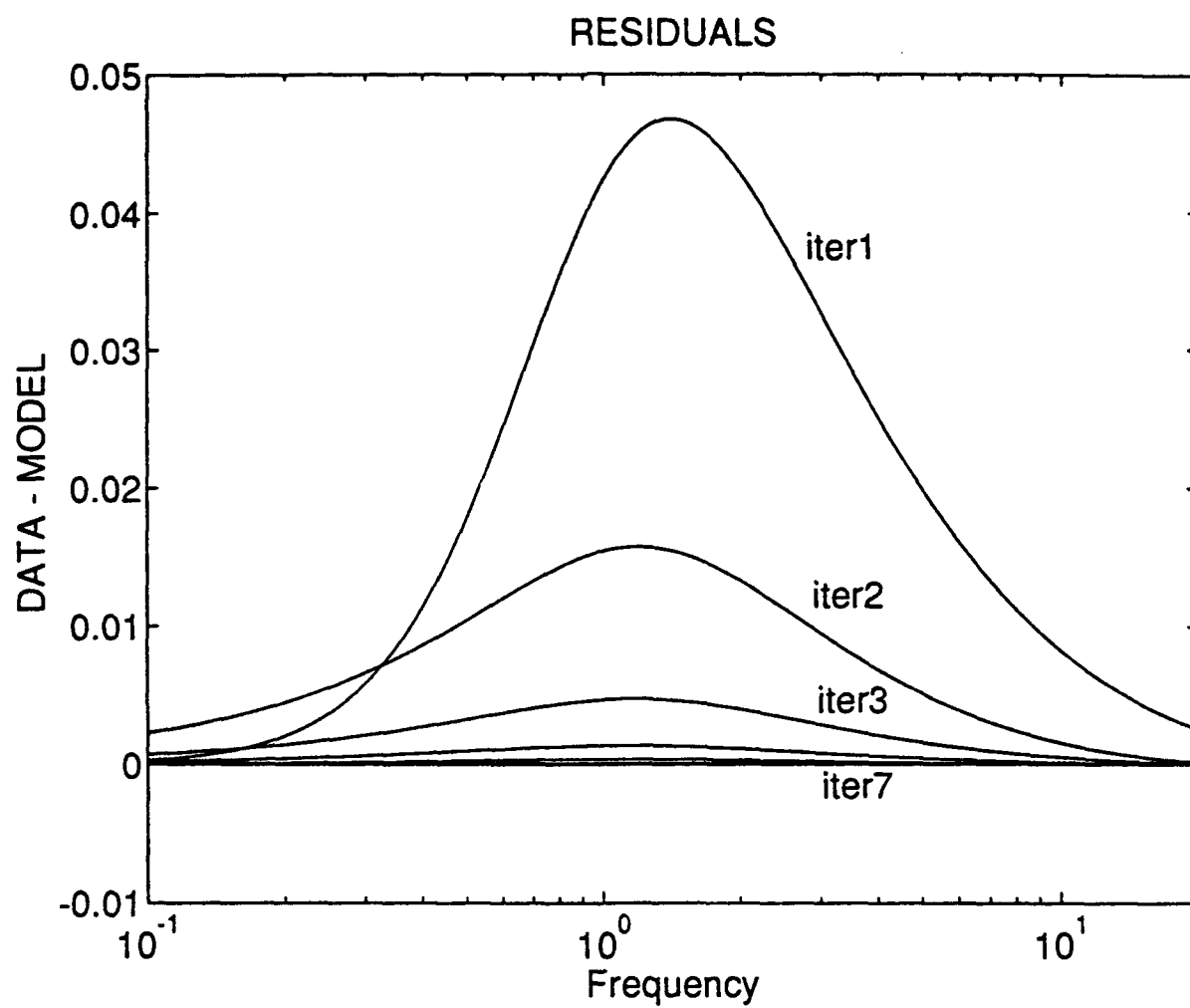


Figure 3.23 Residual change by iteration. After a few iterations, prewhitening function approaches to the true model.

parameters and to investigate the degree of bias introduced by overshoot in steady state RDP, Brune's model was applied to the data with overshoot. Data were generated from the VSB model with parameters of $\Psi_{\infty}=1$, corner frequency=1.0 and $Q=50$. Overshoot(B) ranged from 1 to 11. Uniformly distributed random noise with a mean of 1 was multiplied by each corresponding amplitude to simulate perturbation of amplitudes in the logarithmic scale. Brune's model was used as a forward model. The modified scheme was used for an inversion with damping parameter(zero to 300). The Hanning window was applied to get full damping effect. Inversion was performed from five different initial values. The Brune's model introduces a consistent bias(Figure 3.24) in some parameters(steady state RDP and corner frequency) due to the misfit between the forward model and the data. Steady state RDP and corner frequency show increasing bias with increasing overshoot. The effect on steady state RDP is almost linear with overshoot, while the corner frequency approaches an asymptotic value larger than the real corner frequency. Asymptotic behavior of the corner frequency seems to be related to the definition of corner frequency in each model(Appendix A). Note that f_{app} is different from f_{BR} in Appendix A since f_{app} is the result of inversion which takes the least square error from the data even though both represent corner frequency in the Brune's model. Apparent corner frequency, therefore, takes the value between $\sqrt{2} \cdot f_{VSB}$

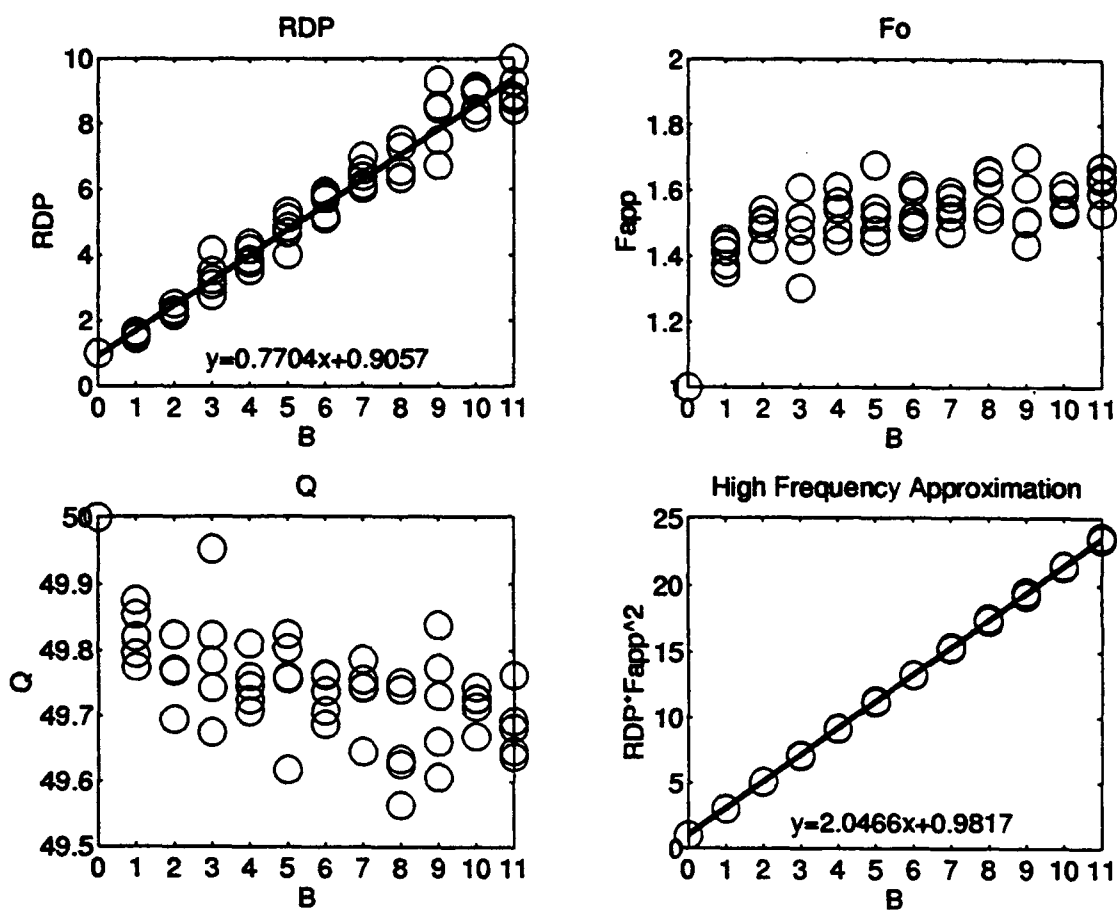


Figure 3.24 Bias introduced by applying the Brune's model to the data with various overshoots. Regression lines are included.

and $f_{BR} (\sqrt{(2B+1)}*f_{VSB})$ if we have sufficient data below corner frequency. Since this is not the case of real application in which low frequency information is limited to a certain amount, there is a trend that the f_{app} is extended to only a slightly larger value than $\sqrt{2}*f_{VSB}$. Therefore, Asymptotic value($f_{app}=1.6*f_{VSB}$) of apparent corner frequency in this empirical test is the result of characteristic of least-square inversion. Apparent corner frequency approaches to the asymptotic value rapidly also as is expected from Equation A-4. Therefore, apparent corner frequency is applicable to estimating real corner frequency(f_{VSB}) if overshoot exists.

Asymptotic behavior of the corner frequency with respect to overshoot determines the degree of bias in steady state RDP since high frequency approximation is always stable. High frequency approximation is independent on the size of explosion since there is a scaling relation between steady state RDP and corner frequency as was introduced in Chapter 1. High frequency approximation plot in Figure 3.24 illustrates above relation in another way. In Figure 3.24 with constant steady state RDP and corner frequency, the value of steady state RDP times square of corner frequency increases with a constant rate as B increases. The slope and intercept in the regression line represents apparently $(2B+1)$ which is the effect of overshoot to the spectral shape in the von Seggern-Blanford model. Therefore,

$$(\Psi_{\infty \text{app}}) * (f_{\text{app}}^2) = (\Psi_{\infty \text{true}}) * (f_{\text{VSB}}^2) * (2B_{\text{true}} + 1)$$

If we put $f_{\text{app}}^2 = 1.6^2 * f_{\text{VSB}}^2 = 2.56 * f_{\text{VSB}}^2$, then

$$\Psi_{\infty \text{app}} = (0.78 * B_{\text{true}} + 0.385) * (\Psi_{\infty \text{true}}).$$

The slope of the above equation is close to the slope obtained by the linear regression in steady state RDP plot (Figure 3.24). Above approximation cannot be applied if B is less than one since $f_{\text{app}} \neq 1.6 * f_{\text{VSB}}$ and induce discrepancy of the intercept between the above equation and the regression line.

Above relations between estimated and real parameters make it possible to apply the Brune's model as a forward model even where overshoot exists. Even though overshoot biases the corner frequency and steady state RDP, their biases are so consistent that the real parameters can be derived. Q is not sensitive to overshoot and shows consistent accuracy within a range of test if we admit the perturbation of Q by the noise.

Inversion with the VSB model

When the VSB model is applied to the ideal data ($\Psi_{\infty}=1$, $B=2$, $f_0=1.5$ and $Q=50$) contaminated by theoretical noise, the schemes without damping hardly resolve the parameters with the data including overshoot effect. When Scheme 1 is used,

none of the 30 randomized initial inputs converge with the relaxation parameter of 0.5. This results from the high degree of dependence between parameters (Figure 3.25), especially between overshoot and Q . Figure 3.25 was obtained by dropping the relaxation parameter to 0.2 and selecting initial parameters cautiously. When the initial input was selected carefully, the outputs is not far from the expected values. Figure 3.26 is the result of inversion with Scheme 2. Scheme 2 shows high correlation between parameters, too. The condition number ranges from 10^5 to 10^3 in both cases. When the modified scheme was used, correlation coefficient (Figure 3.27), resolution matrices (Figure 3.28) and condition number (Figure 3.29) are all acceptable. Figure 3.30 illustrates one of the matches between the data and expected model. The histogram for the estimates from 30 calculations show the mean values of parameters ($\Psi_\infty=0.97$, $B=1.34$, $F_o=1.78$ and $Q=50.5$). When compared with the expected parameters ($\Psi_\infty=1$, $B=2$, $F_o=1.5$ and $Q=50$), they show some discrepancies in overshoot. The bias of an overshoot is related to the noise at low frequencies. The noise can affect in estimating either steady state RDP or overshoot or both. A slight bias of corner frequency is the result of the trade-off with overshoot (Equation A-4). In most cases when the data have either the theoretical or background noise, it is dangerous for a limited bandwidth data set to apply an overshoot estimation unless it is large. Even though it can

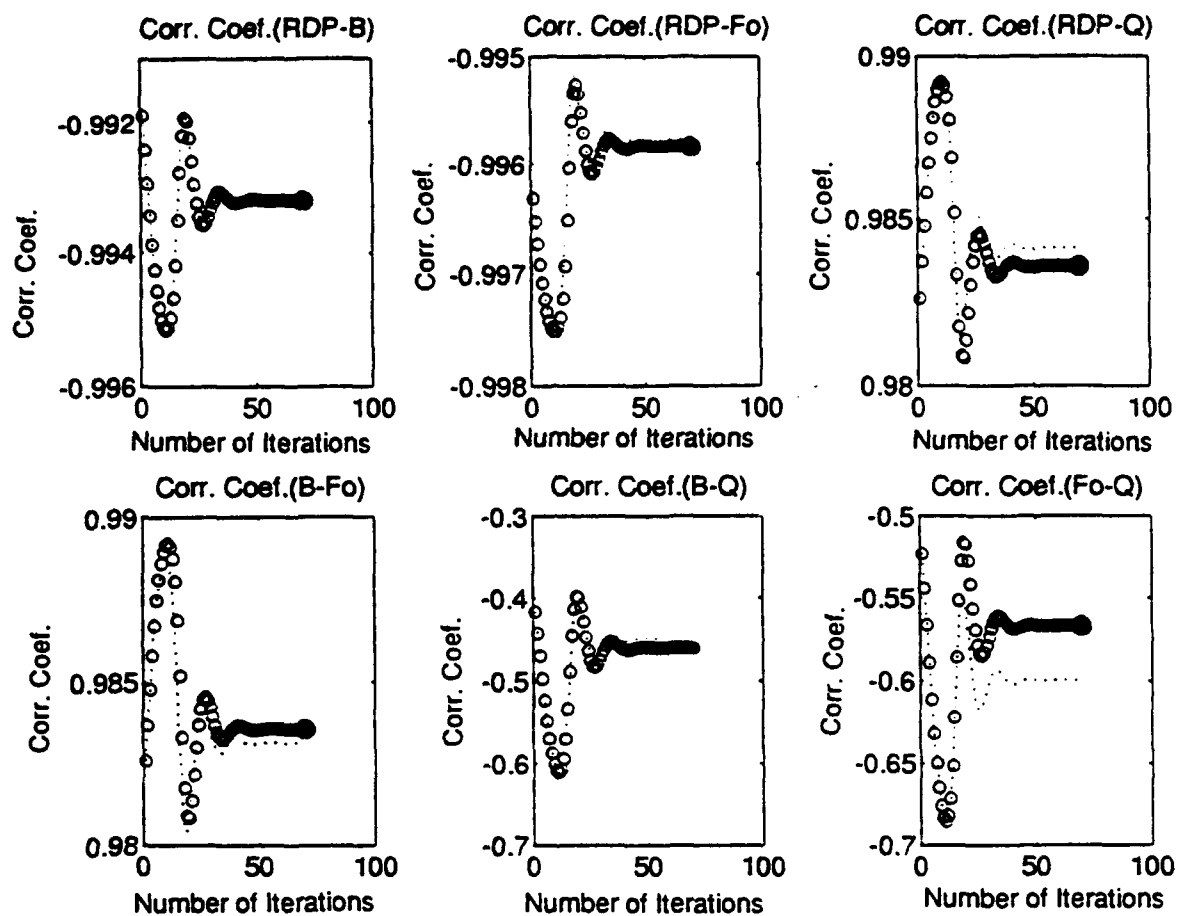


Figure 3.25 Correlation coefficients between parameters when the VSB model and Scheme 1 are used. Parameters show large correlations.

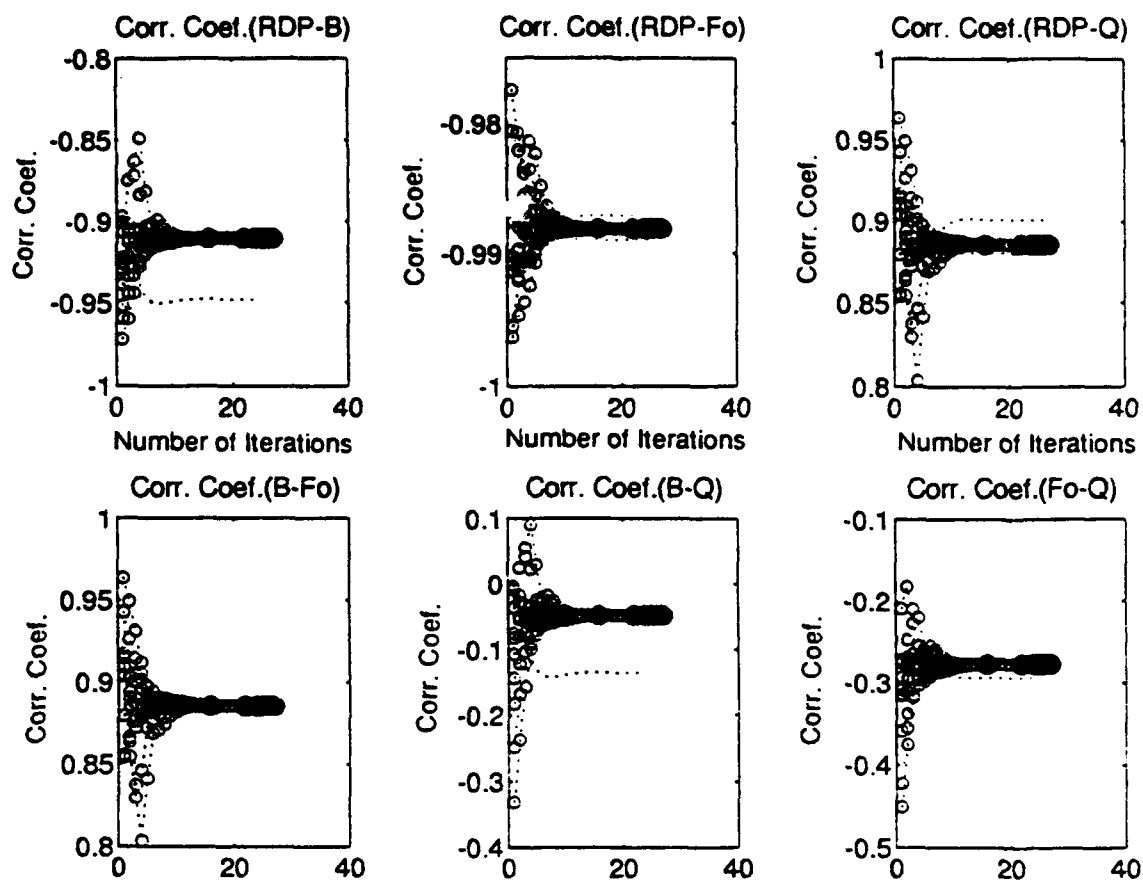


Figure 3.26 Correlation coefficients between parameters when the VSB model and Scheme 2 are used. Parameters show large correlations.

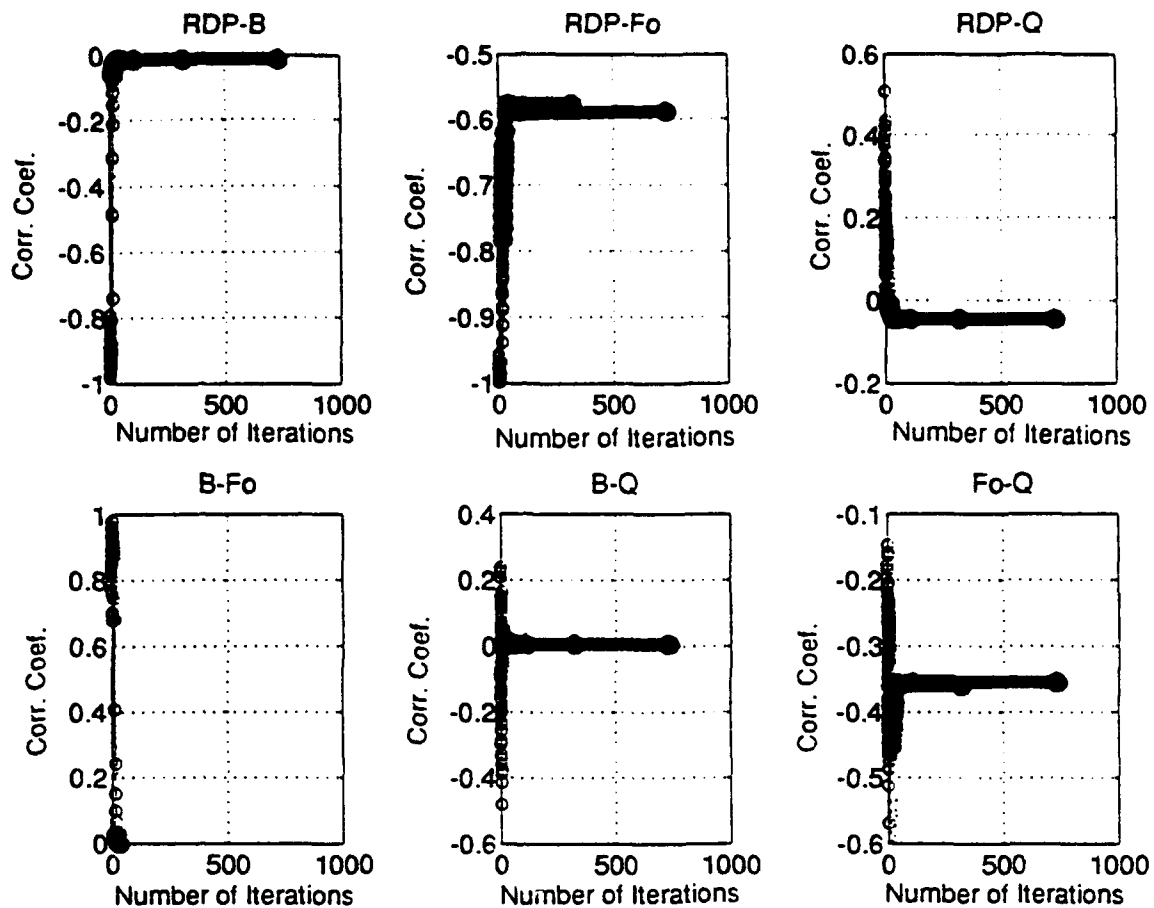


Figure 3.27 Correlation coefficients between parameters when the VSB model and the modified scheme are used. Parameters show improved correlation coefficients.

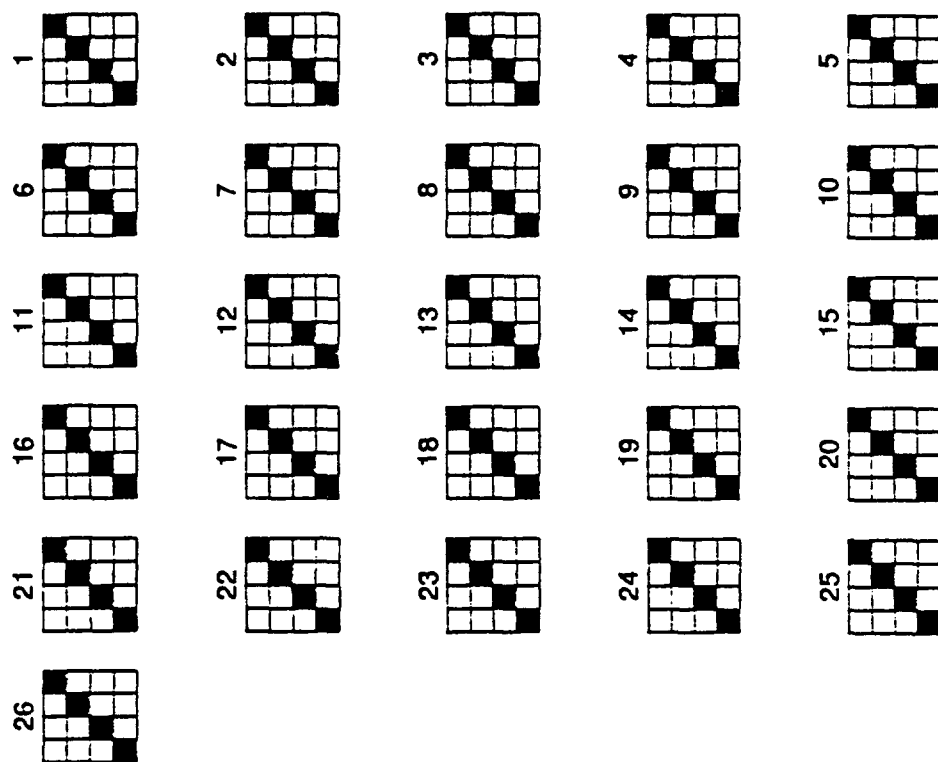


Figure 3.28 Resolution matrices for the modified scheme when the VSB model was applied to the data with overshoot. Meshes represent matrices in an order of Ψ_∞ , B , f_0 and Q in both rows and columns. The matrices is shaded in 64 levels of gray(0=white). The number on the right of each matrix denotes iteration number.

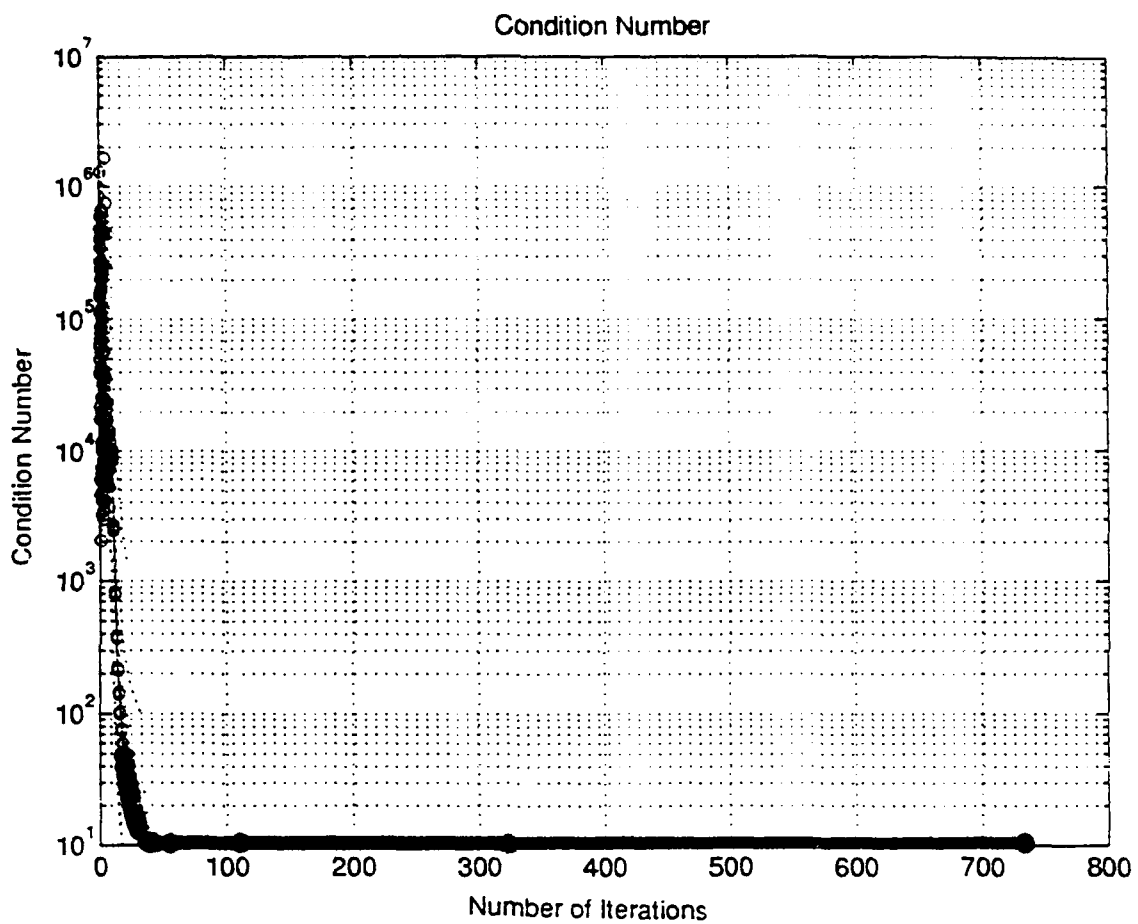


Figure 3.29 Condition numbers in the modified scheme.
The VSB model is applied to the data with overshoot.

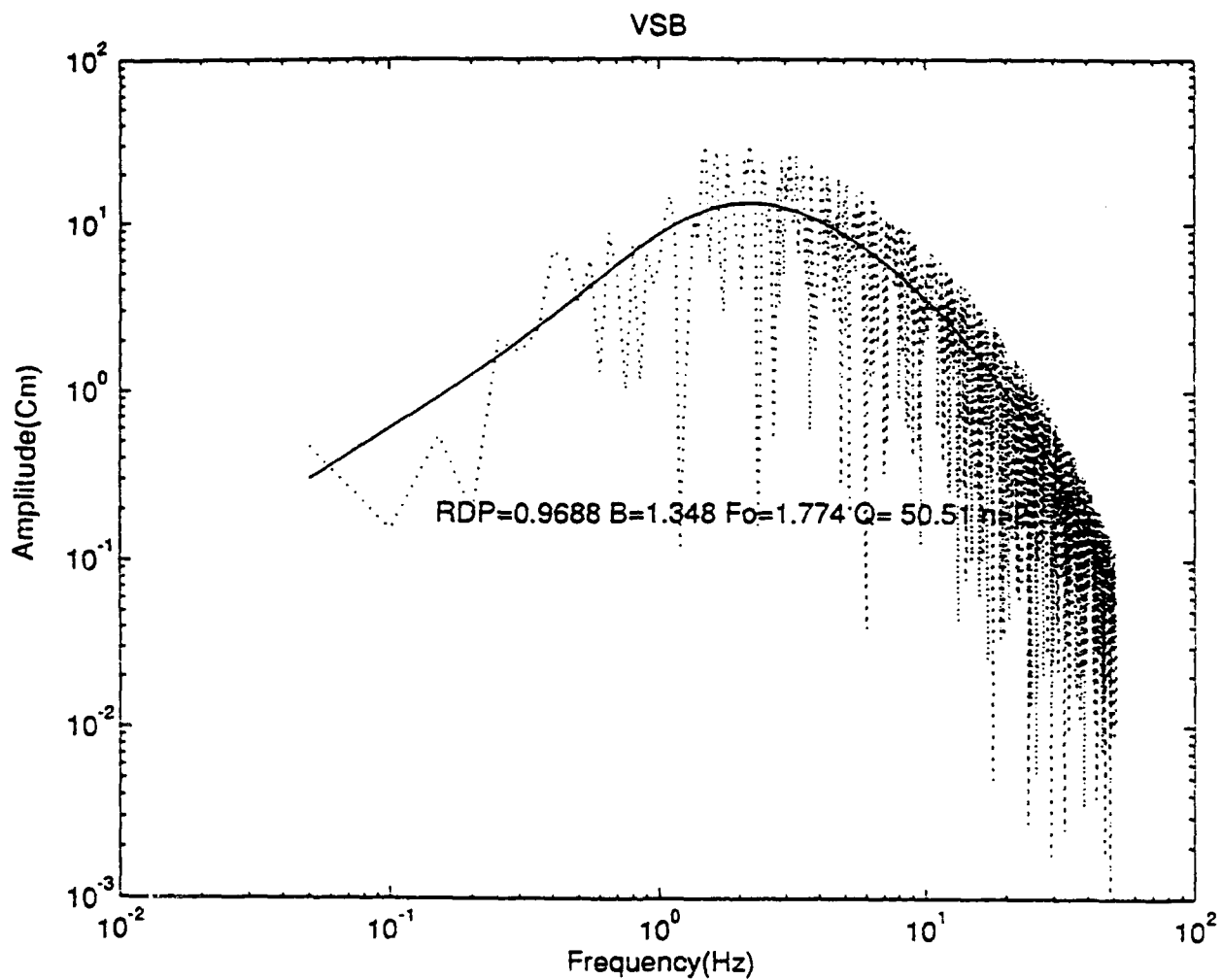


Figure 3.30 One of the inversions performed by the VSB model with the modified scheme. Resolved parameters are illustrated in the middle of the figure.

be resolved maintaining high independence on the other parameters by inserting large damping parameter, the variation due to noise may exceeds the resolution of the overshoot.

Near-Field Effect

In order to examine the effect of the near-field propagation term to the estimation of source parameters using the simple far-field propagation model, data were generated at source-receiver distances of 0.1 km to 5 km at 0.2 km intervals (< 1 km) and 1 km and the interval of 1 km intervals at greater distances. The source parameters used were $\Psi_{\infty}=1$, $B=2$, $F_0=1.5$ and $Q=50$. P-wave velocity was 1 km/sec. Synthetic velocity spectra including the near-field term were generated with the VSB model (Equation 2.7). Theoretical random noise was added to the ideal data. The near-field and far-field contributions for an explosion source are discussed in Appendix D. The modified inversion scheme was used with maximum damping value of 200.

Figure 3.31 shows the estimated parameters with respect to distance. Inversion was performed with several different initial values in all cases. As shown in the figure, near-field effect biases the estimation of source parameters at short distances. Since the near-field effect is strongest at low frequencies, the bias of steady state RDP and overshoot are largest. It is meaningless to treat steady state RDP and

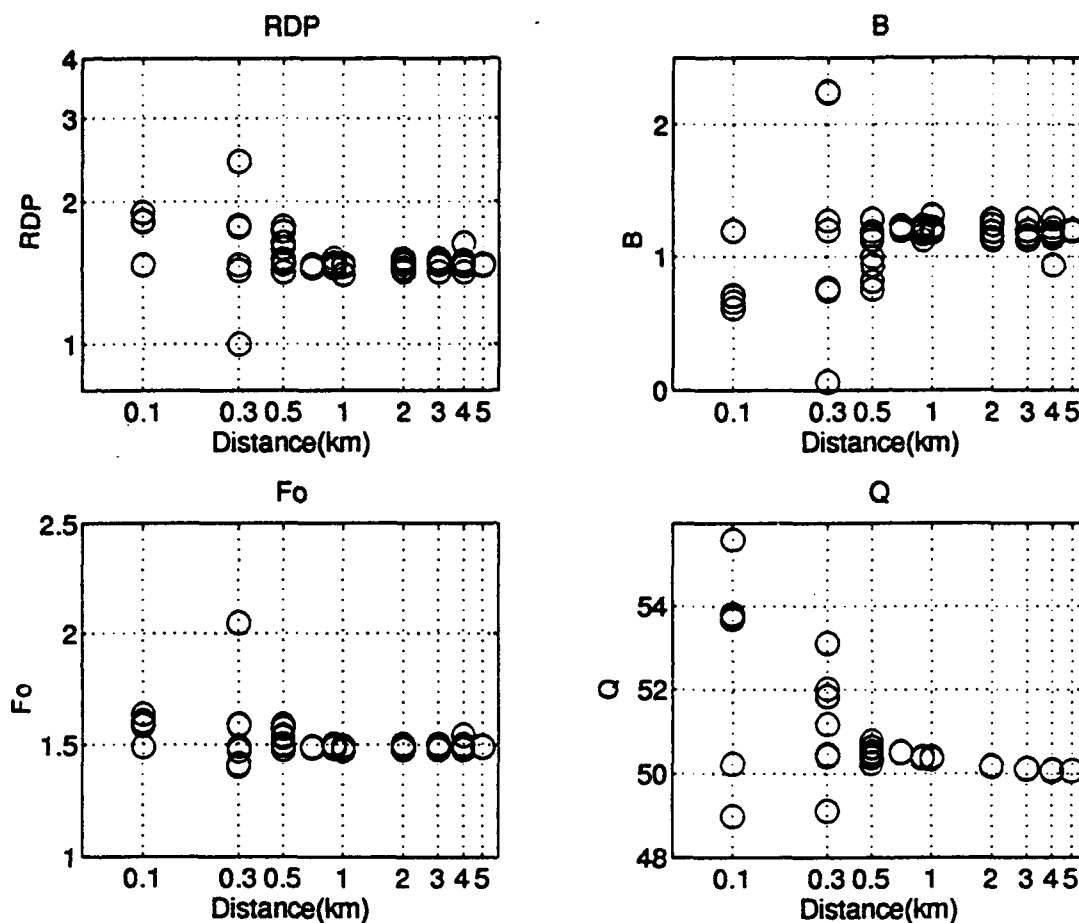


Figure 3.31 Parameter estimation with far-field source model to the data including near-field term. Modified scheme was used. Steady state RDP and overshoot are biased(Expected value; $\Psi_{\infty}=1$, $B=2$). Corner frequency and Q show reliable estimation(Expected value; $f_0=1.5$ Hz, $Q=50$).

overshoot as a separate parameter in short ranges when applying far-field source model. These values approach asymptotically to the true values as the distance increases and as the near-field effect diminishes, but they still show some bias. This bias at large distances is also related to the noise level as was mentioned in the previous section. On the other hand, corner frequency shows stable result beyond 0.5 km in this test. Corner frequency bias is a result of trade-off with overshoot, rather than of the near-field effect. Q shows considerably stable result throughout the entire range of interest since it is dependent only on the slope of high-frequency decay which is unaffected by the near-field term. The high-frequency approximation, $(\Psi_{\infty}) \cdot (2B+1) \cdot f_0^2$, is stable (Figure 3.32) but slightly smaller than the expected value of 11.25.

Inversion of Synthetic Seismogram

Homogeneous Half-Space Synthetic Seismogram

The simplest propagation path correction that can be applied to the data is that for a homogeneous full-space. In order to explore the applicability of full-space path model in the near-source region, synthetic seismogram in a homogeneous half-space were generated and the source parameters were estimated by an inversion. For the generation of the Green's function, the Cagniard-de Hoop

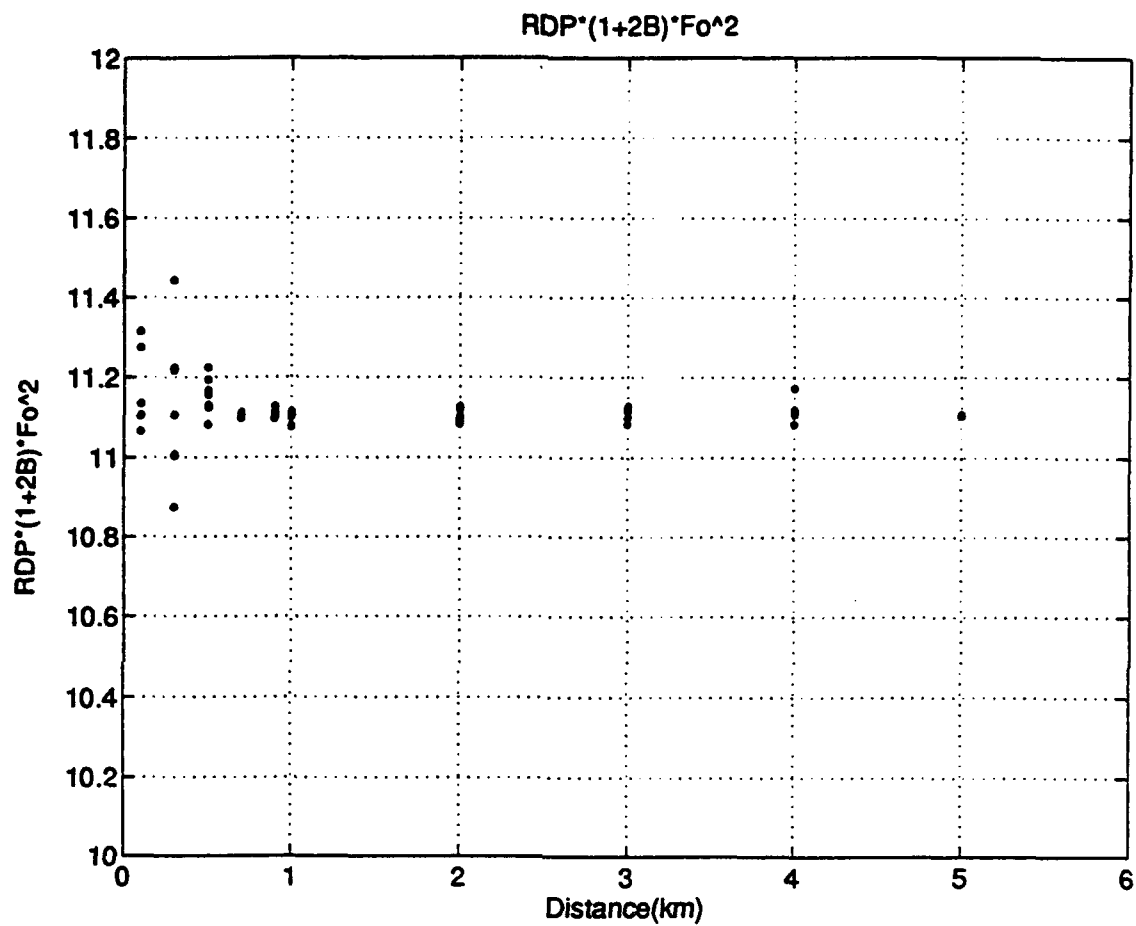


Figure 3.32 High frequency approximations almost constant although there are some perturbations at very short distances.

method(Johnson, 1978) was used with the P-wave velocity of 2.2 km/sec, S-wave velocity of 1.23 km/sec(Poisson's ratio=0.27) and the density of 1.85 gram/cm³. These are the physical properties of the material near the source point of the Rainier Mesa nuclear explosions investigated later. The depth of the source was taken as 0.395 km. The source function of von Seggern-Blanford was convolved with the Green's function to produce the synthetic seismograms. Attenuation was not applied to generate the synthetic seismogram. Figure 3.33 shows the radial and vertical component of synthetic velocity seismogram with source parameters of $B=2$ and $f_0=1.5$ Hz. The corresponding source strength is 2.8×10^{20} dyne-cm in moment, or 261 m³ of steady state RDP. Epicentral distance ranges from 0.1 km to 5 km. The duration of the seismogram is 8.192 seconds with the sampling rate of 250 samples/sec. One second of time delay was applied to seismograms at distances greater than 2 km. The number written on the right corner of each seismogram in Figure 3.33 is the peak amplitude in cm/sec. Comparing both components, the effect of face waves are greater on the vertical component than on the radial component because of the elliptic motion of the Rayleigh waves. Figure 3.34 shows the maximum velocity amplitudes with respect to source-receiver range. Vertical component(close circle) shows consistent decrease in its amplitude(slope=-1.68) with

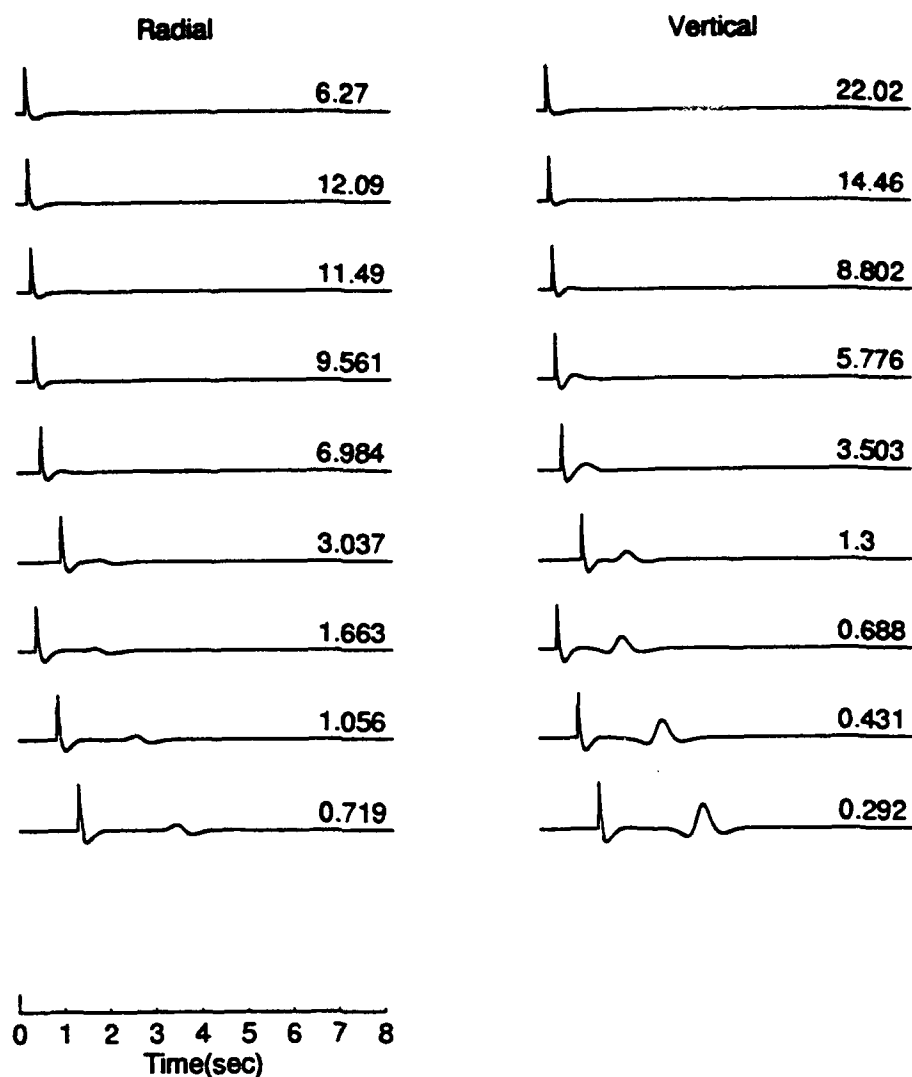


Figure 3.33 Radial and vertical velocities derived from the homogeneous half-space model. The ranges are [0.1, 0.3, 0.5, 0.7, 1.0, 2.0, 3.0, 4.0, 5.0] km from the top. One second of time delay was applied beyond 1 km. Maximum amplitudes are illustrated on the left top of each seismogram.

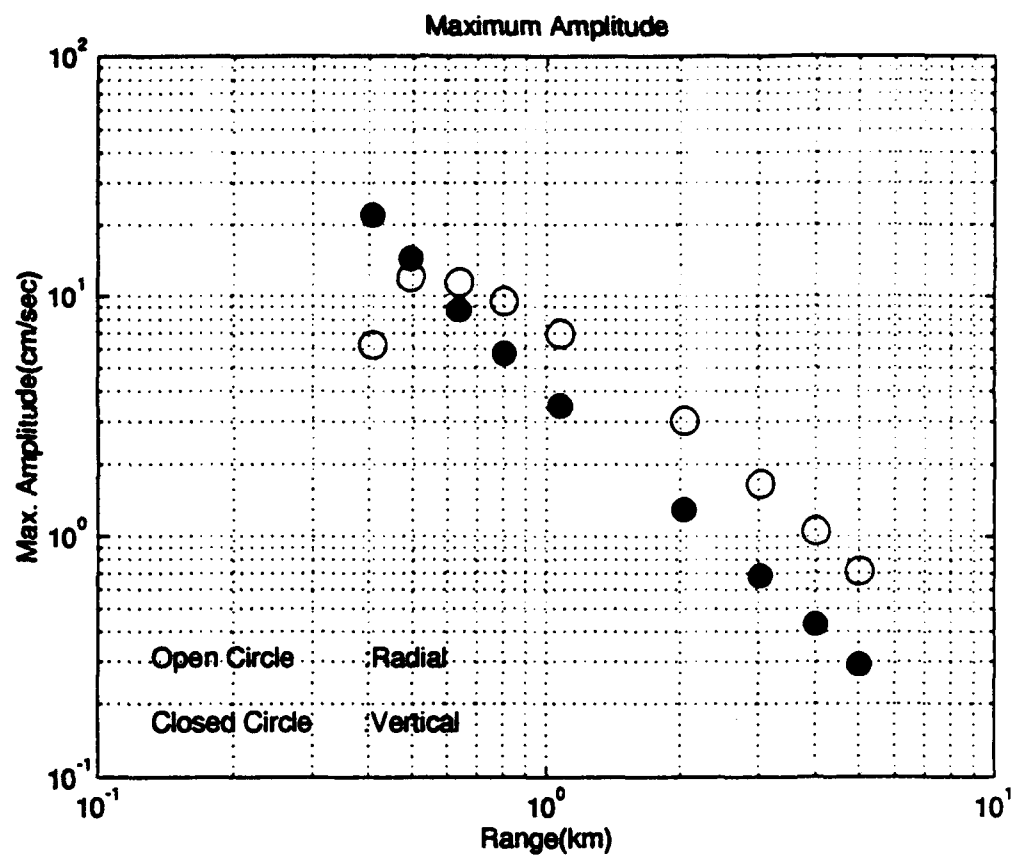


Figure 3.34 Maximum amplitude distribution with respect to distance. Ranges are calculated from the straight paths. Open circles denote radial components and closed circles denote vertical components. Geometric effect is shown clearly.

respect to distance while radial component (open circle) does not by the geometric effect.

Inversion of homogeneous half-space data with full-space path model

Data obtained by Cagniard-de Hoop method in the homogeneous half-space and convolved with the VSB source model were inverted using the homogeneous full-space far-field VSB source model. Q is fixed at 10000 which is large enough not to affect in high frequencies. The results of inversion are plotted in Figure 3.35 and 3.36. Figure 3.35 is the result of radial components while Figure 3.36 is the result of vertical components. For each calculation, ten different initial values were used. At first glance, the estimated parameters are not consistent with respect to distance. They also show non-unique behavior beyond 2km for the vertical component. Steady state RDP and overshoot show different estimates in both components from different initial inputs. The large variances of steady state RDP beyond 2 km for the vertical data are related to the trade-off with overshoot. The effect of surface waves at low frequencies and the interference between body waves and surface waves (Shift Theorem) can make the slope steep at low frequencies before the corner frequency. The steep slope before the corner frequency increases the overshoot estimate to infinity which cannot be met in a numerical calculation

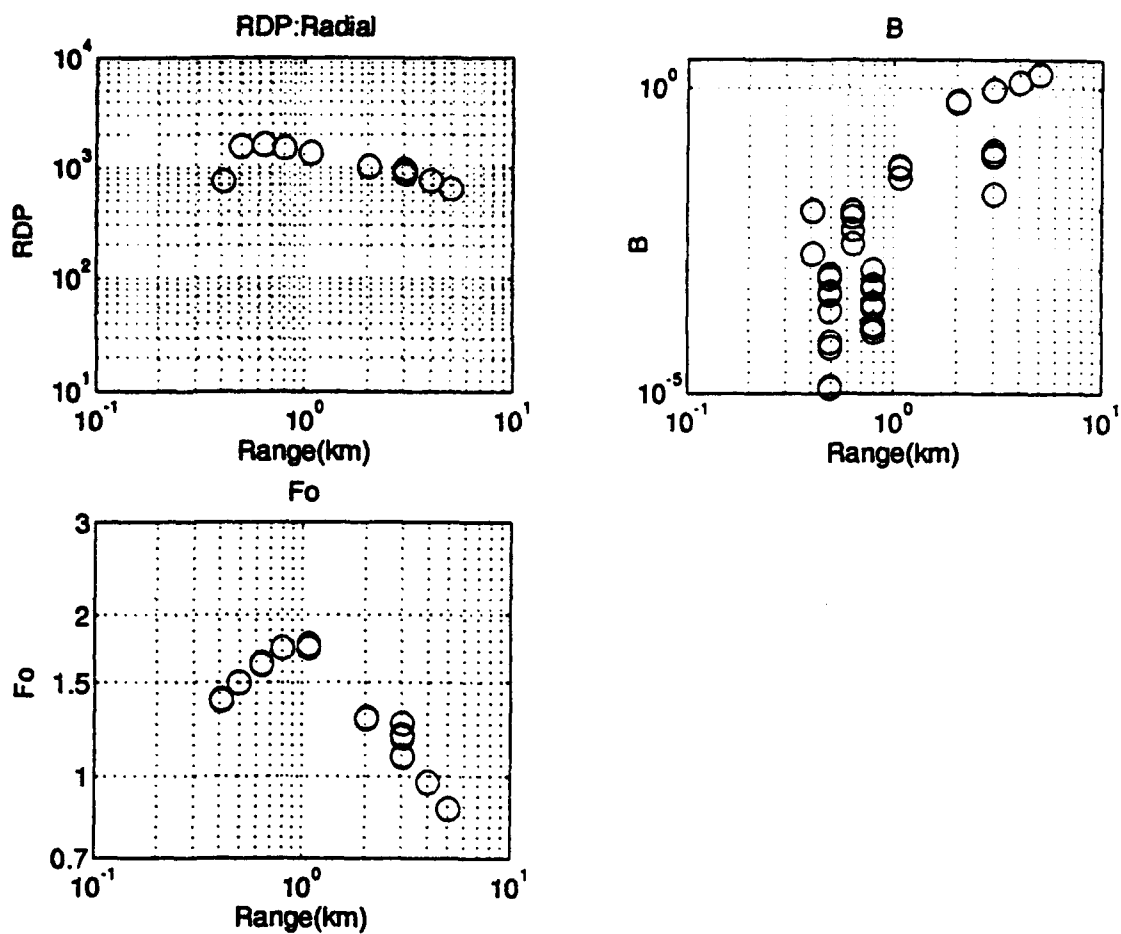


Figure 3.35 Parameter estimation from the radial components. The far-field VSB model was used. Q was not applied.

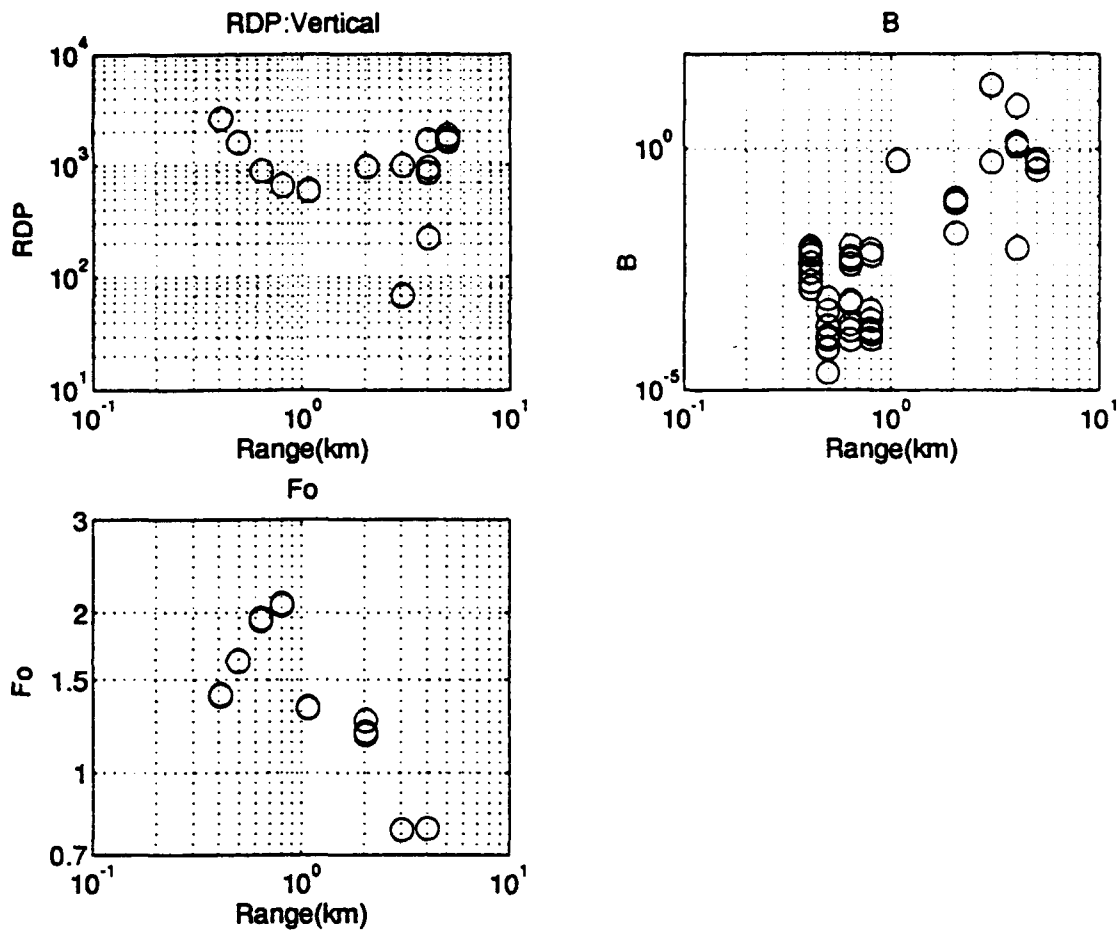


Figure 3.36 Parameter estimation from the vertical components. The far-field VSB model was used. Q was not applied.

and the iteration will be ended by the termination criteria. Since each inversion with different initial values has different convergence rate, the iteration generally is ended at different values of parameters before the iteration arrives at the minimum in the least-square sense. It should be noted that the inversion never converges to the termination criteria without a damping parameter. In this case, overshoot grows larger and larger while steady state RDP drops farther and farther as a trade-off. If steady state RDP drops to within the limit of computer round-off error, the change of steady state RDP between iterations can be on the order of the error, and the rate of change can be quite large. In this case, the iteration proceeds forever. The surface wave effect is more serious in the vertical component than in the radial component because of the elliptic characteristic of the Rayleigh waves in a homogeneous half-space. The flat slope at low frequencies where the near-field term is important cannot be explained by the far-field forward model with positive overshoot. In this case, since the overshoot cannot be negative because of constraint, the inversion approximates to as zero forever. This is the same kind of non-convergent problem as in the far-field data, but steady state RDP shows small variance since the overshoot effect is insignificant. The role of damping parameter in both cases is simply to limit the convergence rate. Large damping parameter reduces convergent

rate and incidentally meets the termination criteria. Therefore, it is not surprising that there is no consistency in Figure 3.35 and 3.36 which result from large damping parameter and different initial values.

This results illustrate that it is not realistic to try to separate overshoot from other source parameters without considering the path effect. True source overshoot cannot be distinguished from the phenomenological overshoot by path effects such as surface waves or near-field terms. Corner frequency changes as distance changes because of the combined effect of the surface wave's corner frequency and the trade-off with other parameters.

Although individual parameter estimation are not consistent with distance, the high-frequencies are shown to be consistent with simple theoretical considerations (Figure 3.37). The high-frequency data plotted against source-receiver distance in degrees matches well with Zoeppritz's amplitude partition curve (Young and Braile, 1976) at the free-surface. Circles denote the radial components and the pluses denote the vertical components in the figure. Solid and dotted lines are theoretical amplitude partitioning curves derived from a plane wave in a homogeneous half-space by Zoeppritz. Each line denotes S- and P-wave amplitude respectively. They correspond to the radial and vertical component in the free surface. Data obtained from the homogeneous half-space Green's functions ranging from 0.1 to

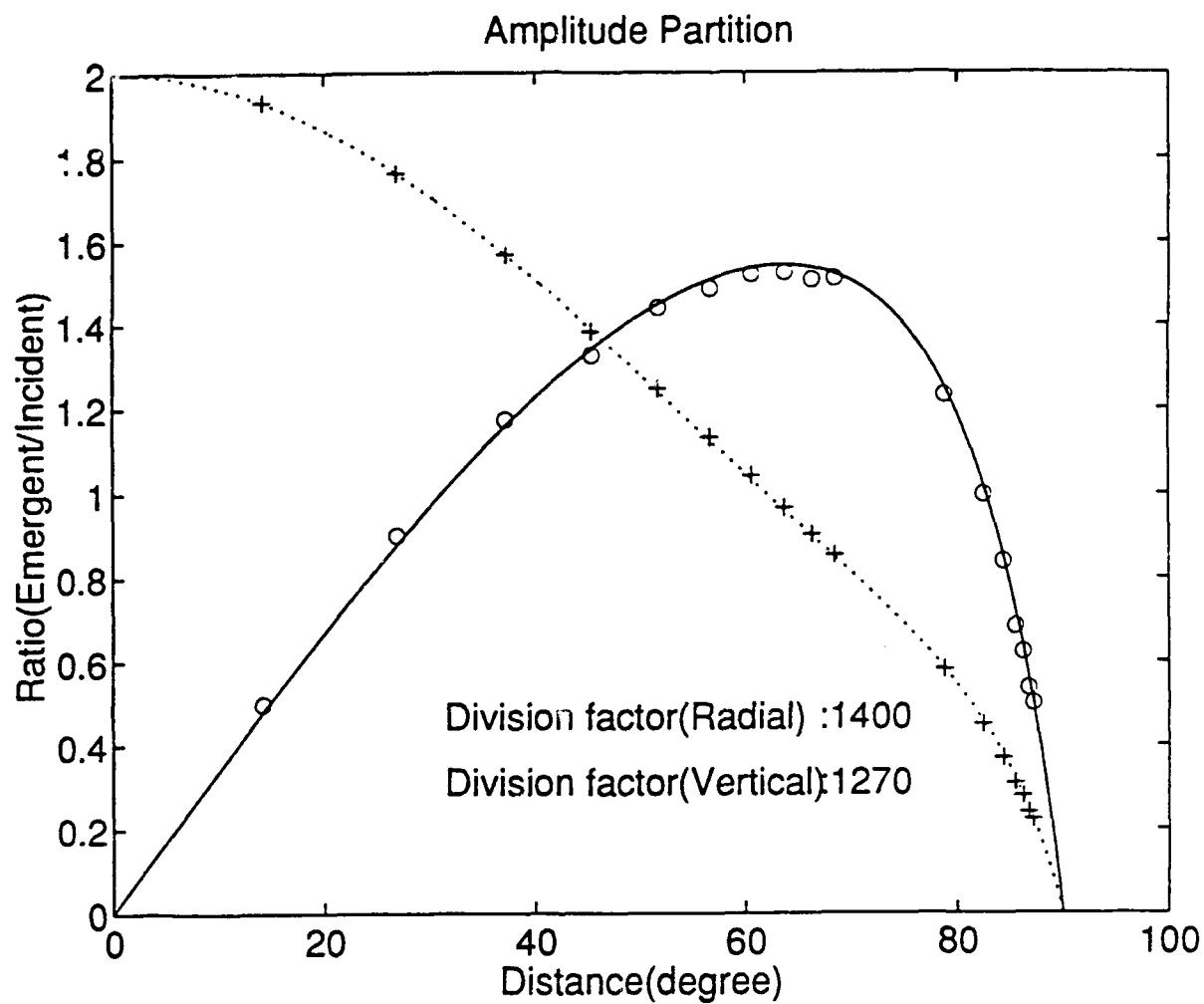


Figure 3.37 High frequency approximation from both components matches well to the Zoeppritz's amplitude partitioning curve.

8 km and the VSB source model with $\Psi_{\infty}=1$, $B=2$ and $f_0=1$ were used in an inversion for the plot. The high-frequency approximations were obtained by multiplying each Ψ_{∞} , $(2B+1)$, and f_0^2 from the result of inversion. Since there is no separation of parameters in the high-frequency approximation, the expected division factor for the amplitude ratio to the input source should be

$$(\Psi_{\infty}) * (2B+1) * (F_0^2) = 261 * 5 * 1^2 = 1305.$$

The curve fit by eye shows that the division factor is 1270 for the vertical component and 1400 for the radial component. This implies that the high-frequency approximation is consistent with distance regardless of the existence of the near-field and the surface wave effect, but that separation of each source parameter cannot be done without considering these effects. It also verifies that the total value of steady state RDP, overshoot and corner frequency can be determined by information at high frequencies because of the large number of data points in this frequency band and that the separation of each parameter is performed based on the information at low frequencies. Combined effect of steady state RDP, overshoot and corner frequency is not affected by the near-field and surface wave effect since both effects are generally confined to low-frequencies.

Near-field consideration

It is almost impossible to separate the near-field effect accurately in the seismogram. On the other hand, it is relatively simple to introduce the near-field term into the simplified theoretical path model. Equation 2-6 or 2-7 are examples of near-field path models in a homogeneous full-space. Homogeneous half-space synthetic seismograms (Figure 3.33) were again used as data. The near-field term was added for the path correction. Figure 3.38 and 3.39 are the results of inversion for radial and vertical components. In Figure 3.39, parameters could not be estimated for epicentral distances greater than 2 km because of the continuous interchange of overshoot (increase) and steady state RDP (decrease). Parameters, especially overshoot and steady state RDP, can be resolved better on the radial component at very short distances by considering near-field effect in the path model. The synthetic test shows that the near-field term cannot be neglected in making reliable source parameter estimations if the receiver is close to the source, generally within a few wavelengths. The high frequency approximation shows almost the same result as the previous analysis without near-field consideration. This result is expected since the high frequency approximation is mostly controlled by the data beyond the corner frequency where the near-field or the surface wave contributions are minor.

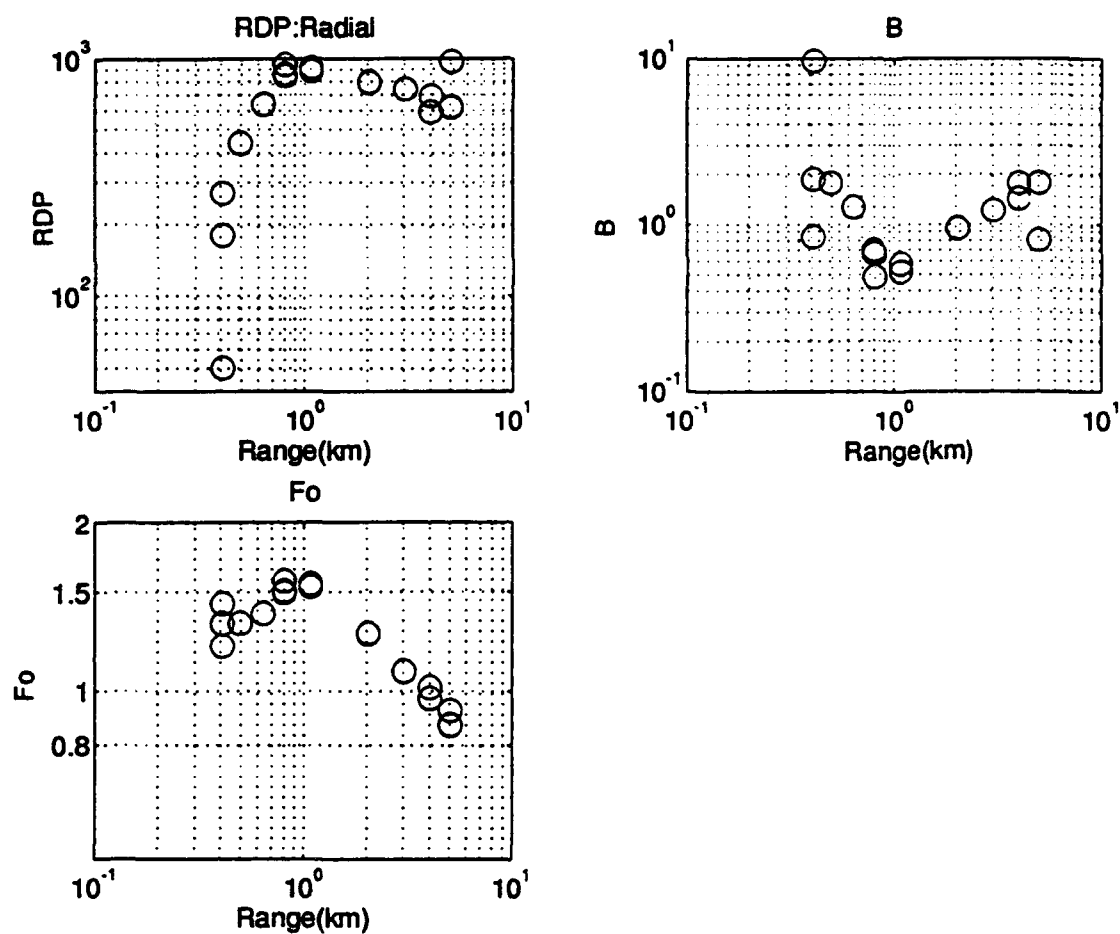


Figure 3.38 Parameter estimation from the adial component. The near-field VSB model is used. Q was not applied.

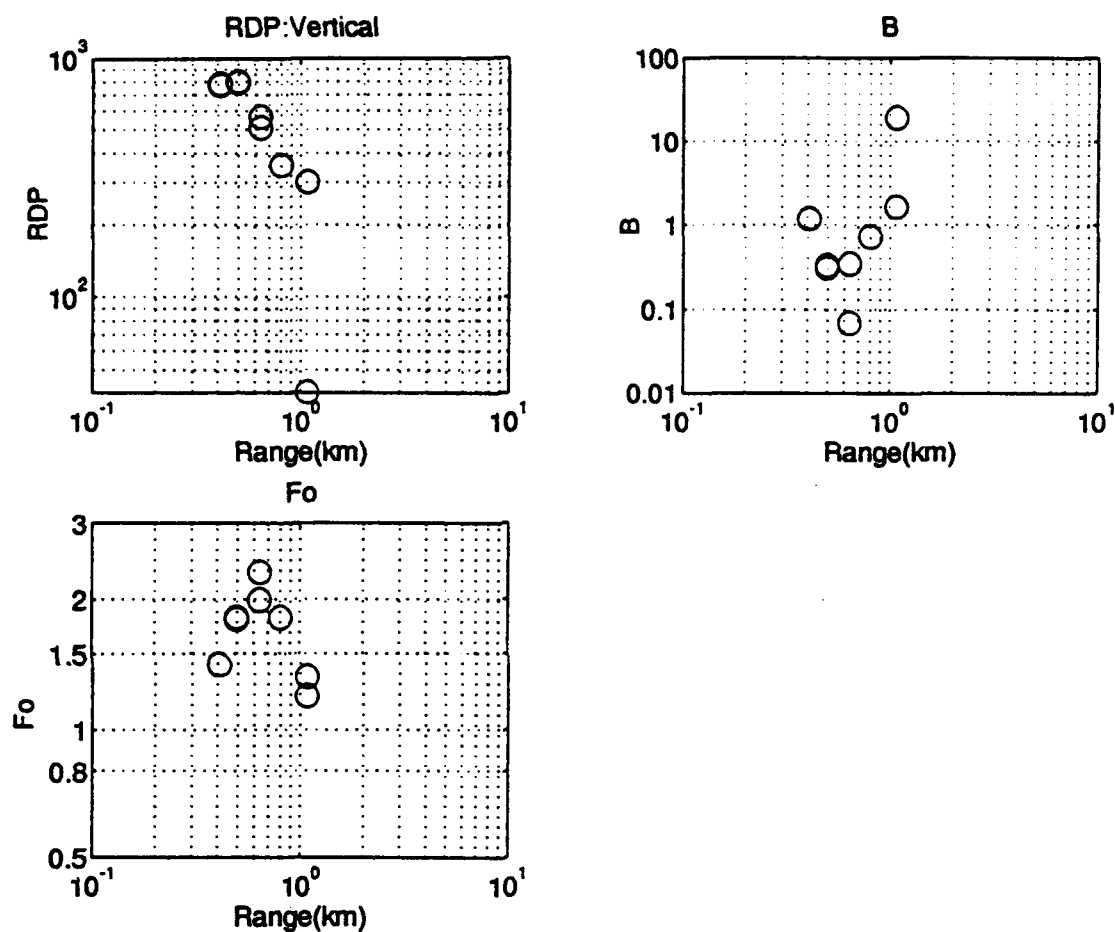


Figure 3.39 Parameter estimation from the vertical component. The near-field VSB model is used. Q was not applied. Beyond 1 km, inversion was not performed since the estimated B tends to increase toward infinity due to the surface wave effect.

Surface wave consideration

Synthetic velocity seismograms in a homogeneous half-space (Figure 3.33) have surface waves which increase in importance with distance. For the separation of surface waves and body waves, it is convenient to use the Green's function if the distance is greater than 1 km because of the difference in body and surface wave phase velocities. Since the Green's function is the response of the media to a delta function and the contribution of surface waves by the Rayleigh poles becomes more and more isolated as the distance increases, the body wave response in the homogeneous half-space shows a sharp peak and is not interfered with the later arrived surface wave response. It is not true if the Green's function is convolved with the source function as in the case of actual observed data. Time separation between two phases is smeared by the much smoother source time function and is much harder to separate each phase than in the original Green's function. The surface wave, in case of Green's function, can be separated from the body wave based by polarization analysis (Figure 3.40). The characteristics of surface waves are shown at 0.5 km, but the amplitude of surface waves starts to exceed the numerical noise at 0.7 km. The surface waves can be separated from the body waves beyond 2 km. The separated phases along with their spectra beyond 2 km are given in Figure 3.41 and 3.42. The analysis was extended to 8 km to check the behavior and the effect of the

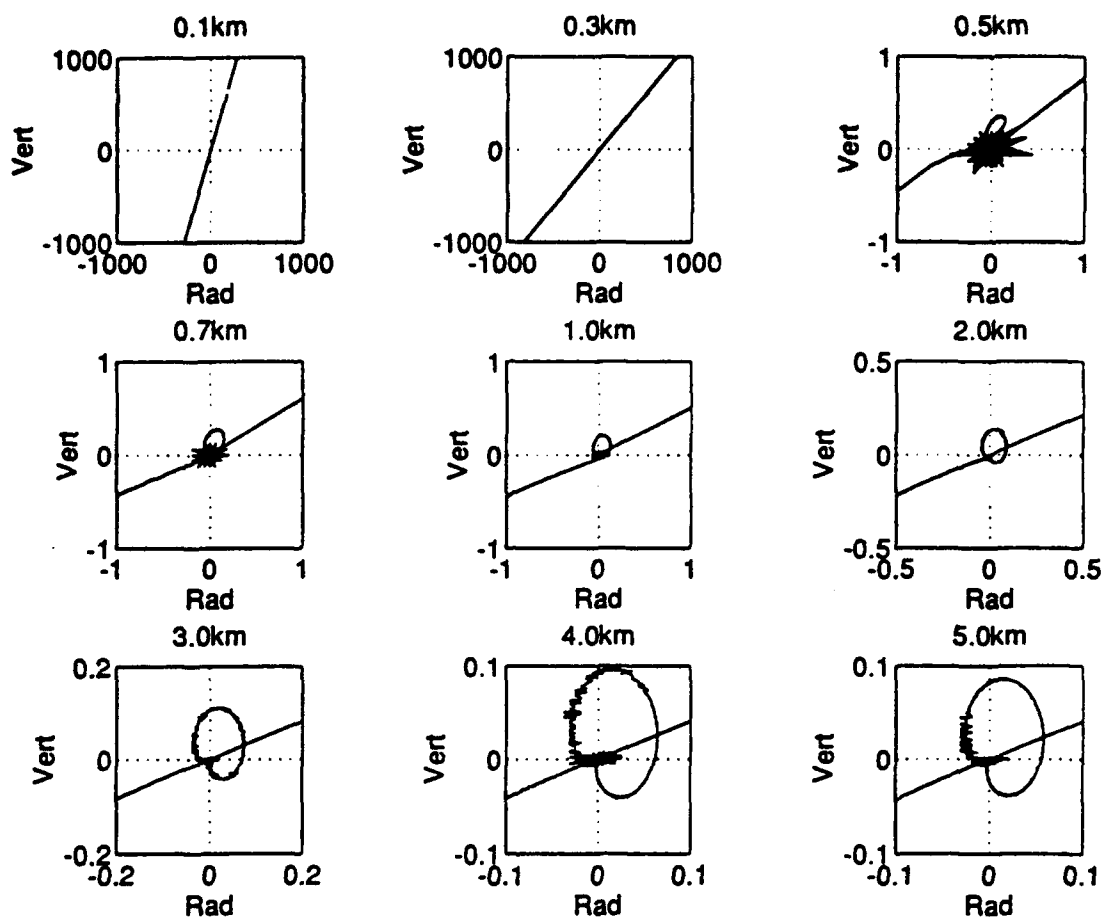


Figure 3.40 Polarization analysis. The amplitude of surface waves starts to exceed the numerical noise at 0.7 km.

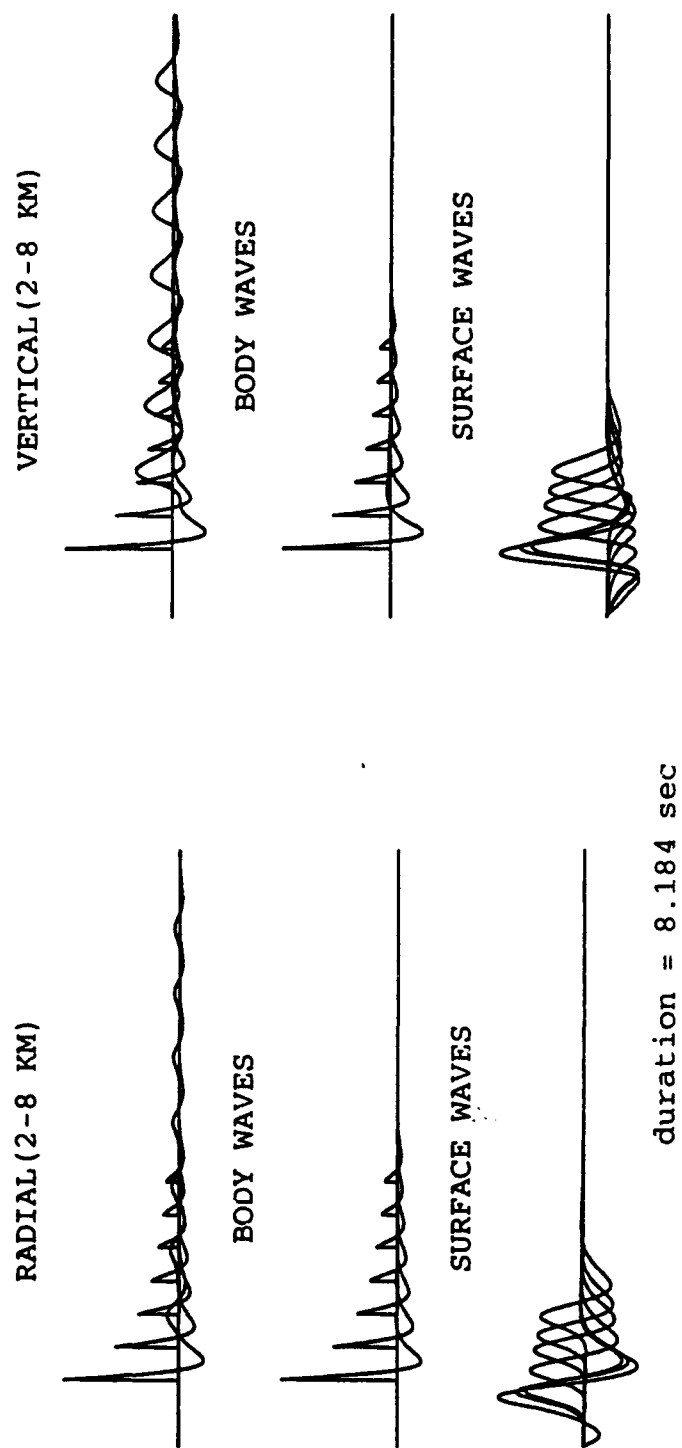


Figure 3.41 Separation of body waves and surface waves from the synthetic seismograms. Epicentral distances range 2 to 8 km.

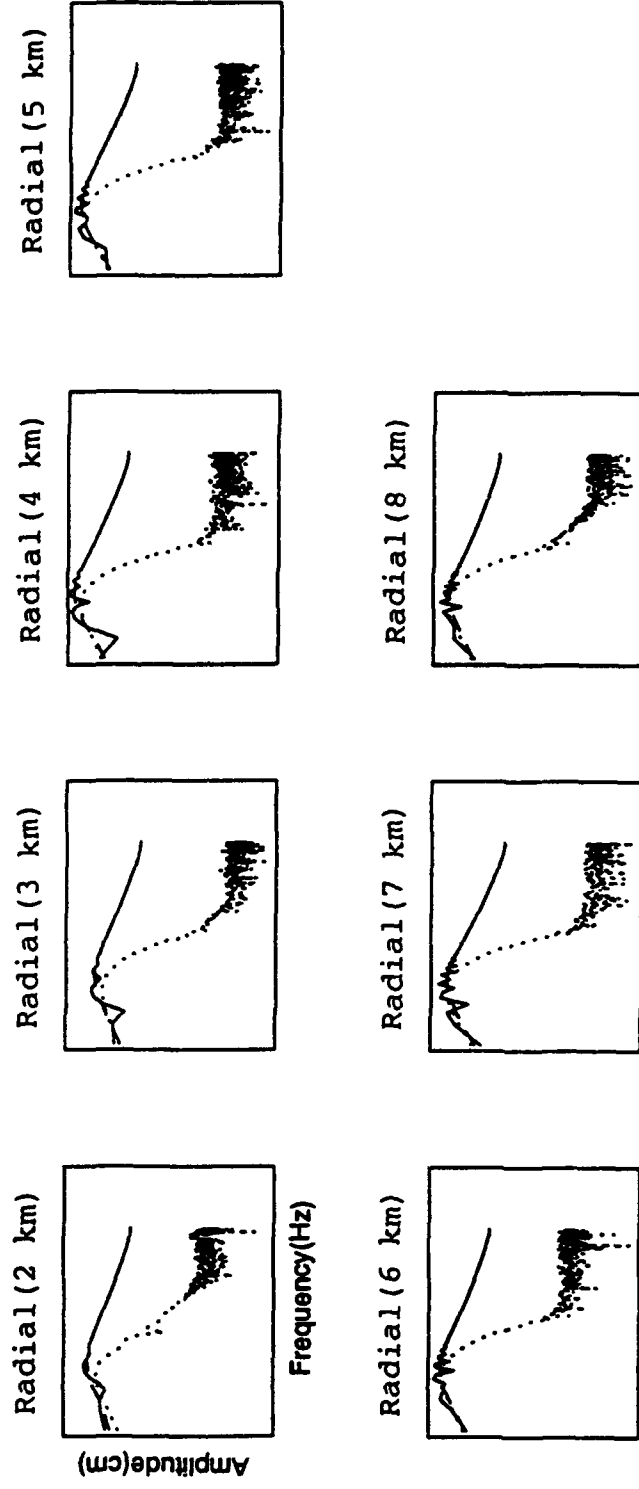


Figure 3.42 Spectra for the synthetic seismogram. Each phase (body wave and surface wave) is separated. The dashed line is the body wave spectrum while the dotted line is the surface wave spectrum. Spectrum of entire wavefield (solid line) is the envelope of both phases.

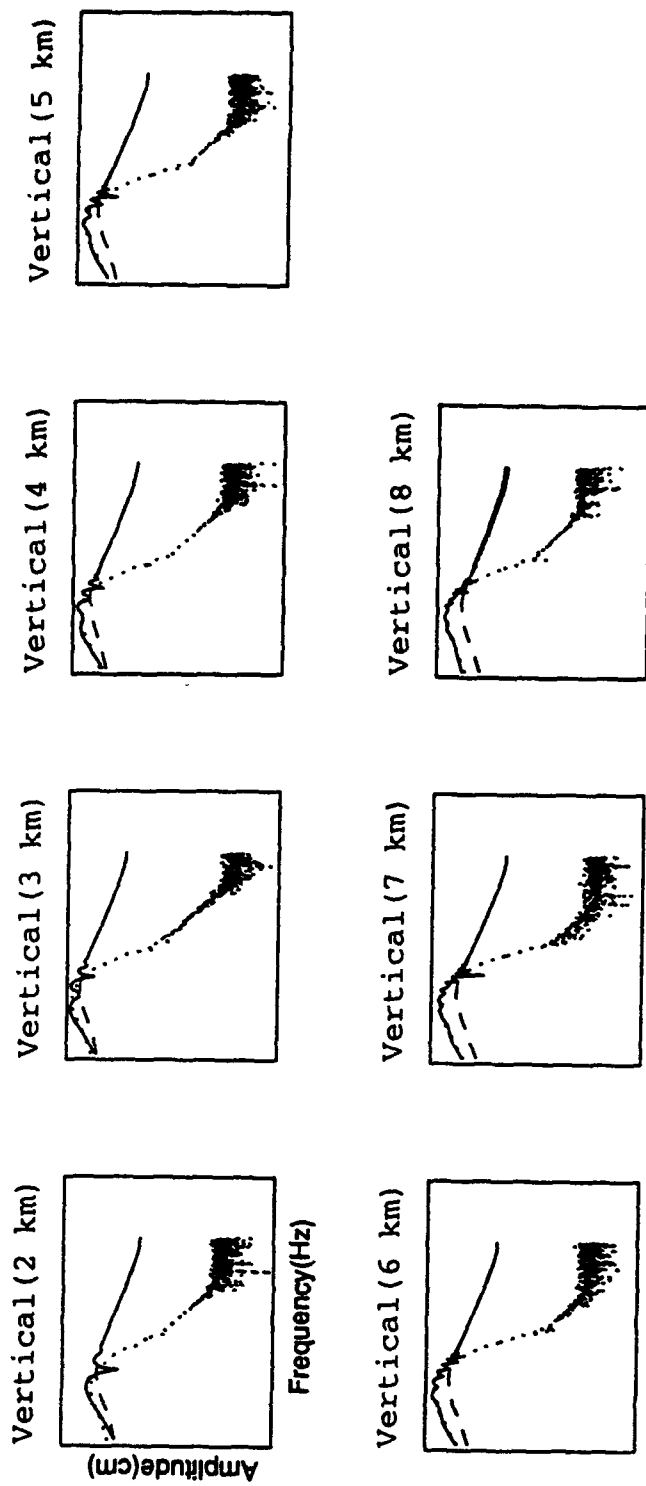


Figure 3.42 Continue

surface waves. In the spectral plots(Figure 3.42), dashed line denotes the body wave spectra while dotted line denotes the surface wave spectra. The entire wave-field spectra(solid line) are the envelopes of both phases. The spectral plot indicates that low frequencies are affected mainly by body and surface waves while the higher frequencies are governed by the body wave only. The spectral plot confirms the stability of high-frequency estimation in the inverse process. For the data generation, the von Seggern-Blanford source time function with parameters of $\Psi_{\infty}=1$, $B=2$, $F_0=1$ Hz were used.

Figure 3.43 is the result of inversion for both components using the body wave portion only. Compared with the estimation of parameters from the entire wave-field in the radial component(Figure 3.44), the parameters from the body wave inversions show smaller variances. In Figure 3.44, circles are the parameters estimated from the entire wavefield while stars denote those of body wave portion. The dashed lines are the means of entire wavefield's parameters and the solid lines are those of body waves. Densely spaced dotted lines are the 90 % confidence interval for the entire wavefield and the sparsely spaced dotted lines are those for the body waves. Since steady state RDP is dependent on the angle of incidence due to the energy partition effect, overshoots and corner frequencies are used for the comparison. In both cases, there are some discrepancies with

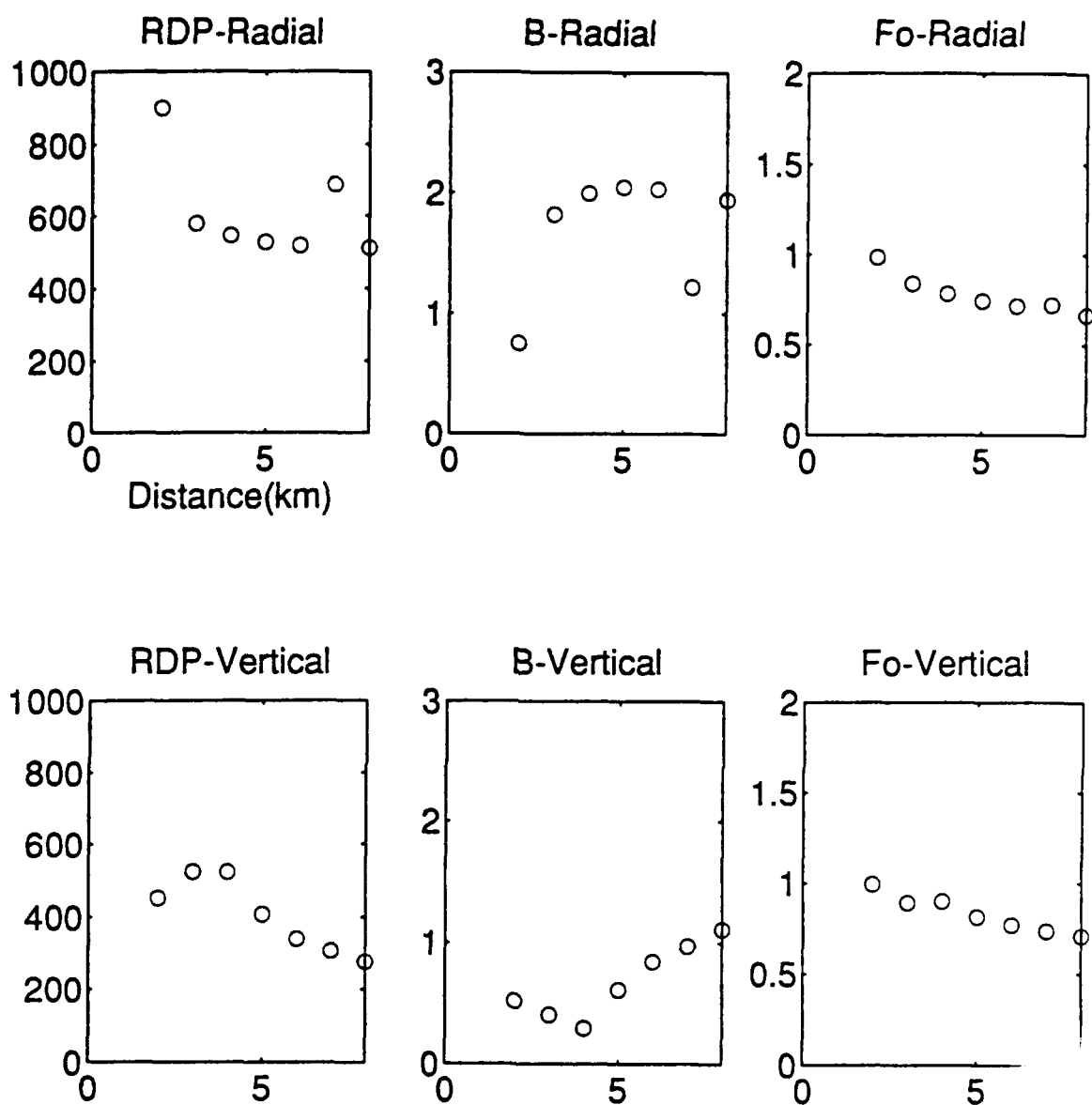


Figure 3.43 Parameter estimation from the body wave portion of the homogeneous half-space synthetic seismograms (2-8 km).

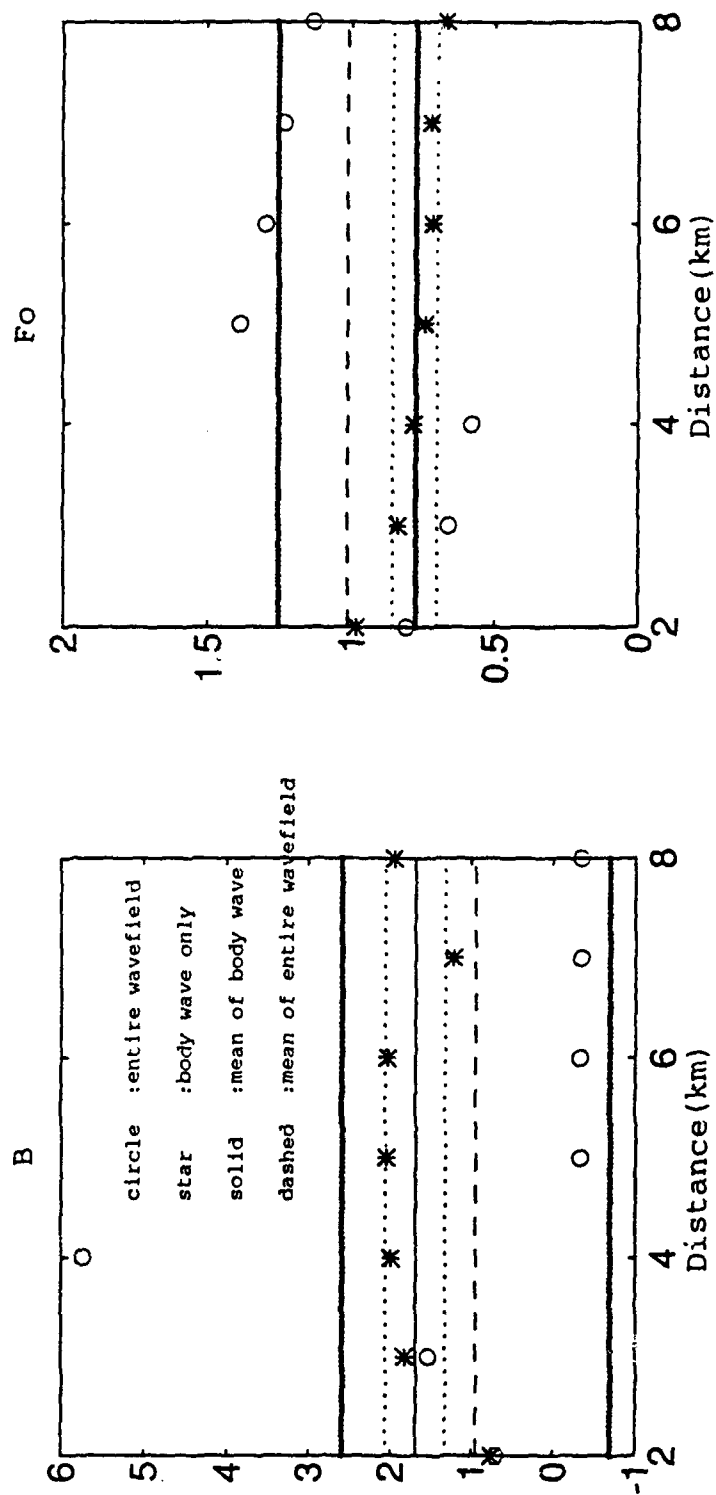


Figure 3.44 Comparison of inversion. When body wave portion is used for an inversion, the estimators show small variances both in overshoot and in corner frequency compared to those of entire wavefield inversion. Both dotted lines are 90 % confidence intervals.

the expected values ($B=2$, $F_0=1$), which is believed to either be due to the effect of the diffracted phases (Gilbert and Laster, 1962) or due to the numerical noises. The comparison of the vertical component shows the biases also. The steady state RDPs are not compared because of the dependence on the incident angle at the free surface.

Layered Structure - Reflectivity Method

Simple tests with homogeneous half-space synthetic seismograms illustrate large biases and trade-offs of the model parameters and the applicability of different source models under certain circumstances. Even though the attenuation effect was not considered in the previous tests, there is a possibility that the high frequency roll-off may be affected by the contribution of surface waves in complex structures. In order to examine this effect more carefully, synthetic seismograms in a layered structure were generated using the reflectivity method (Müller, 1985). The structure for the calculation was designed to match the Rainier Mesa (Olsen et al, 1989). The input parameters are listed in Table 3.6 and the designed structure is shown in Figure 3.45. Thick lines denote the physical properties of the Rainier Mesa from Olsen et al while thin lines are the properties for modelling. Solid and dashed line denote the P- and S-wave velocity respectively while dotted line denotes density structure. S-wave velocities were derived from the P-wave

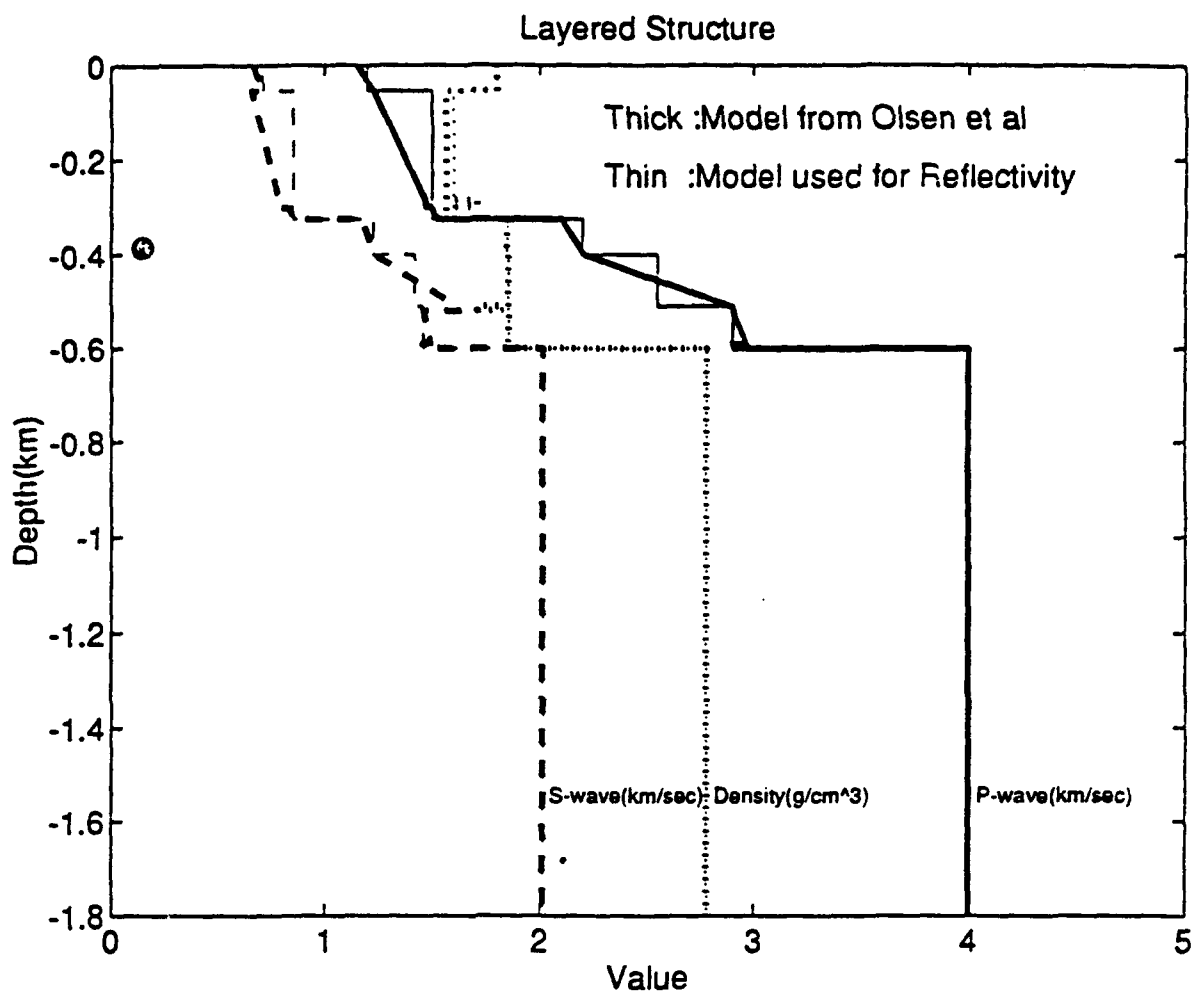


Figure 3.45 Velocity structure at Rainier Mesa. Simplified from the data presented in Olsen et al(1989). Solid line and dashed line denote P-wave and S-wave velocity, while dotted line denotes a density. The location of the source is expressed as a circle.

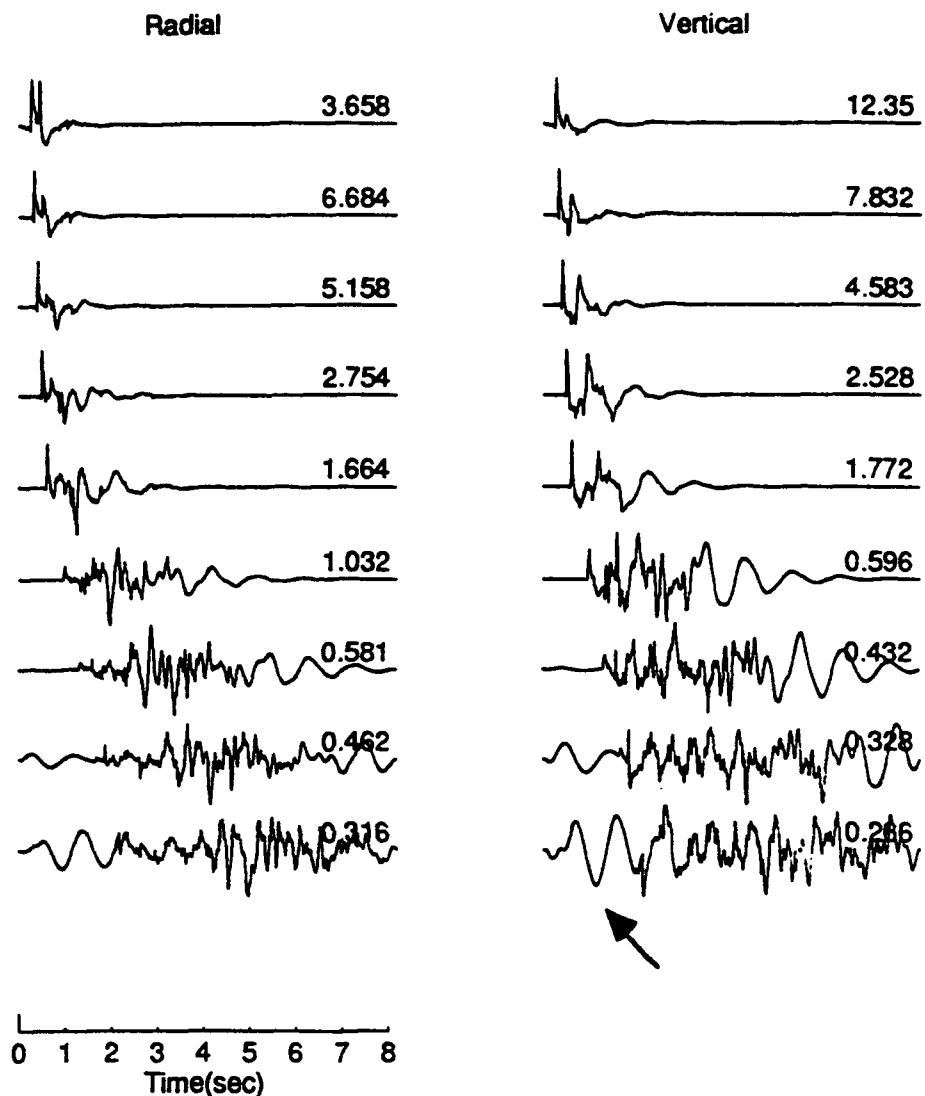


Figure 3.46 Radial and vertical velocities calculated by the reflectivity method and the VSB model ($B=2$, $F_0=1.5$ Hz). The ranges are 0.1, 0.3, 0.5, 0.7, 1.0, 2.0, 3.0, 4.0, 5.0) km from the top. Reduced velocity (8 km/sec) was applied. Maximum amplitudes are illustrated at the right top corner of each esismogram. Wrap around phases (arrow) are shown beyond 4 km.

numbers written on the right corner of each seismogram are the maximum amplitudes in cm/sec. Because of the limited duration, long period wrap around phases, marked as an arrow in Figure 3.46 and believed to be a fundamental mode Rayleigh wave, appear before the first arrival at long distances (4 and 5 km). Figure 3.47 shows ray paths in a given layer and the travel time curve with the Green's function. Ray paths and travel time curves are calculated from the numerical two dimensional ray tracing (Zelt and Smith, 1992) by a Runge-Kutta method and the Green's function is generated by the reflectivity method. At short distances, a converted P-S phase and the surface waves interfere with the direct P-waves, which makes it impossible to separate different phases at this range when the source time function is convolved (Figure 3.48). Figure 3.48 illustrates that most of P-wave energy is concentrated in the direct waves and reflected waves (dashed lines) within 1 km and only in the reflected waves beyond 1 km. Surface wave amplitude exceeds body waves in the range beyond 1 km. All detectable P-waves by the ray tracing are listed in Table 3.6. In Table 3.6, i -angle is the take-off angle in degrees at the source from the horizontal line (clockwise: positive), f -angle is the incident angle to the receiver in degrees, range is the source-receiver distance in km, depth is the depth of the receiver in km, red-time is the reduced time by the reduced velocity ($V_r = 8.0$ km/sec). Code means the type of waves

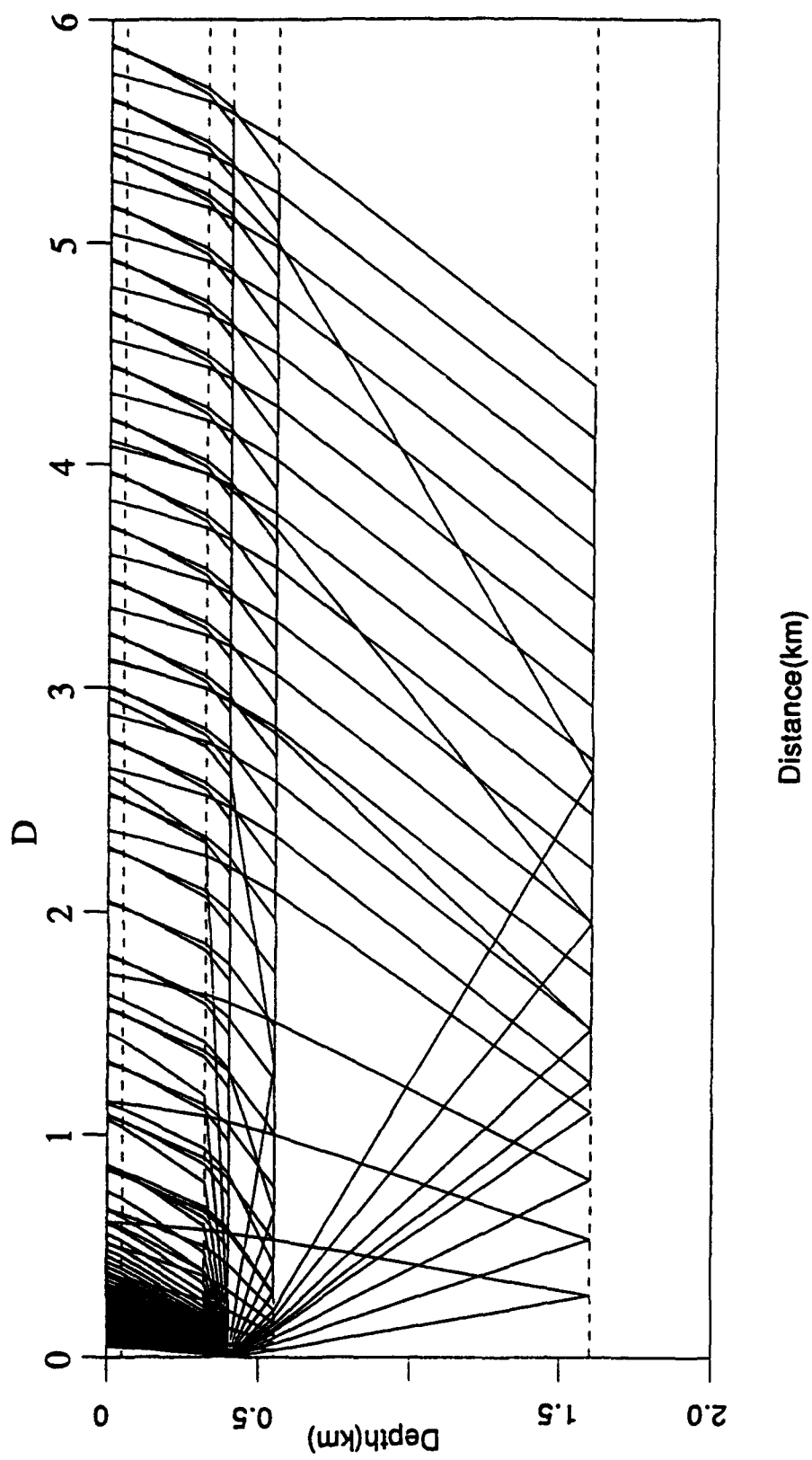


Figure 3.47 Ray paths calculated by the numerical ray tracing(Zelt and Smith, 1992).

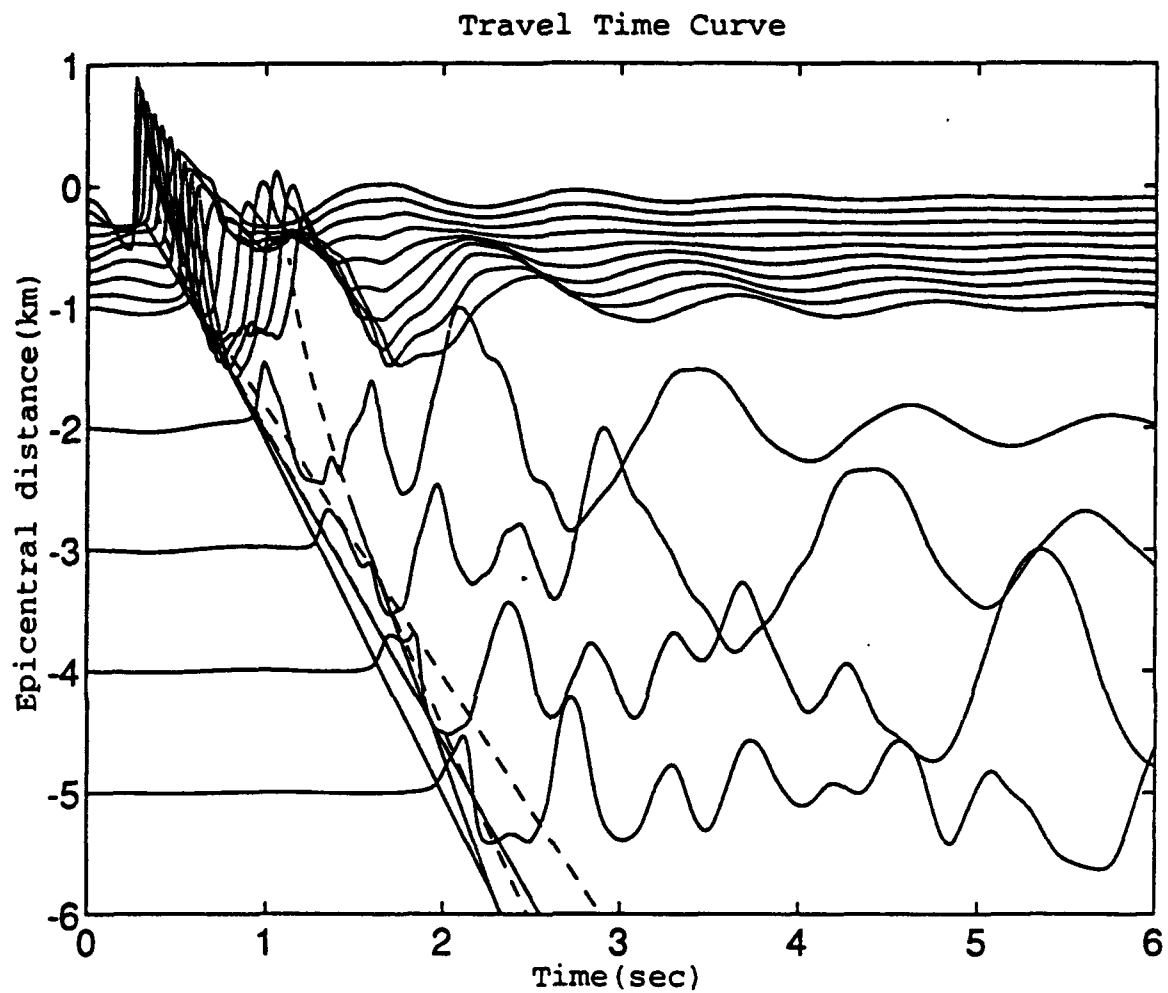


Figure 3.48 Travel time curve. Solid lines are the head waves and the dashed lines are the reflected waves. Vertical component is used.

Table 3.6 Parameters Used in Generating Reflectivity
Seismogram

Depth (km)	α (km/sec)	β (km/sec)	Density (gram/cm ³)	$Q\alpha$	$Q\beta$
0.0000	0.3318	0.0001	0.0013	900(30)	400(17)
0.0000	1.2000	0.9800	1.8000	900(30)	400(17)
0.0050	1.5000	0.8020	1.6000	900(30)	400(17)
0.3250	2.2000	1.2270	1.8460	900(30)	400(17)
0.4000	2.5500	1.4230	1.8520	900(30)	400(17)
0.5100	2.9000	1.4610	1.8520	900(30)	400(17)
1.6000	4.0000	2.0150	2.7800	900(30)	400(17)

*Source depth=0.395 km
 $M_0=1*10^{20}$ dyne-cm
 Type=explosion

(1=refracted waves, 2=reflected waves in its fractional point) at a specific layer(decimal point). For example, if code is 3.2, it represents the reflected waves at layer 3. Note that the numerical method for the ray tracing does not include any S-waves, surface waves and multiples.

Inversion was performed with the synthetic seismograms from the layered structure with small attenuation. Brune's and von Seggern-Blanford's model were used as a forward model. Homogeneous full-space with near-field contribution was assumed as a travel-path of the rays. The P-wave velocity(2.2 km/sec) at the source depth was assumed as an average velocity through the entire ranges. Figure 3.50 and 3.51 show the results of inversion with radial components. In both cases(Brune's and von Seggern-Blanford's model), the estimated parameters cannot be accepted as a reasonable value when they are compared with the expected value($\Psi_{\infty}=180$, $F_0=2.8$ for Brune's model and $\Psi_{\infty}=82$, $B=2$, $F_0=1.5$ for von Seggern-Blanford's model. Q =large in both cases). Expected values for Brune's model are based on the consideration of trade-offs between model parameters as shown in Figure 3.24. For the expected value of steady state RDP, free-surface

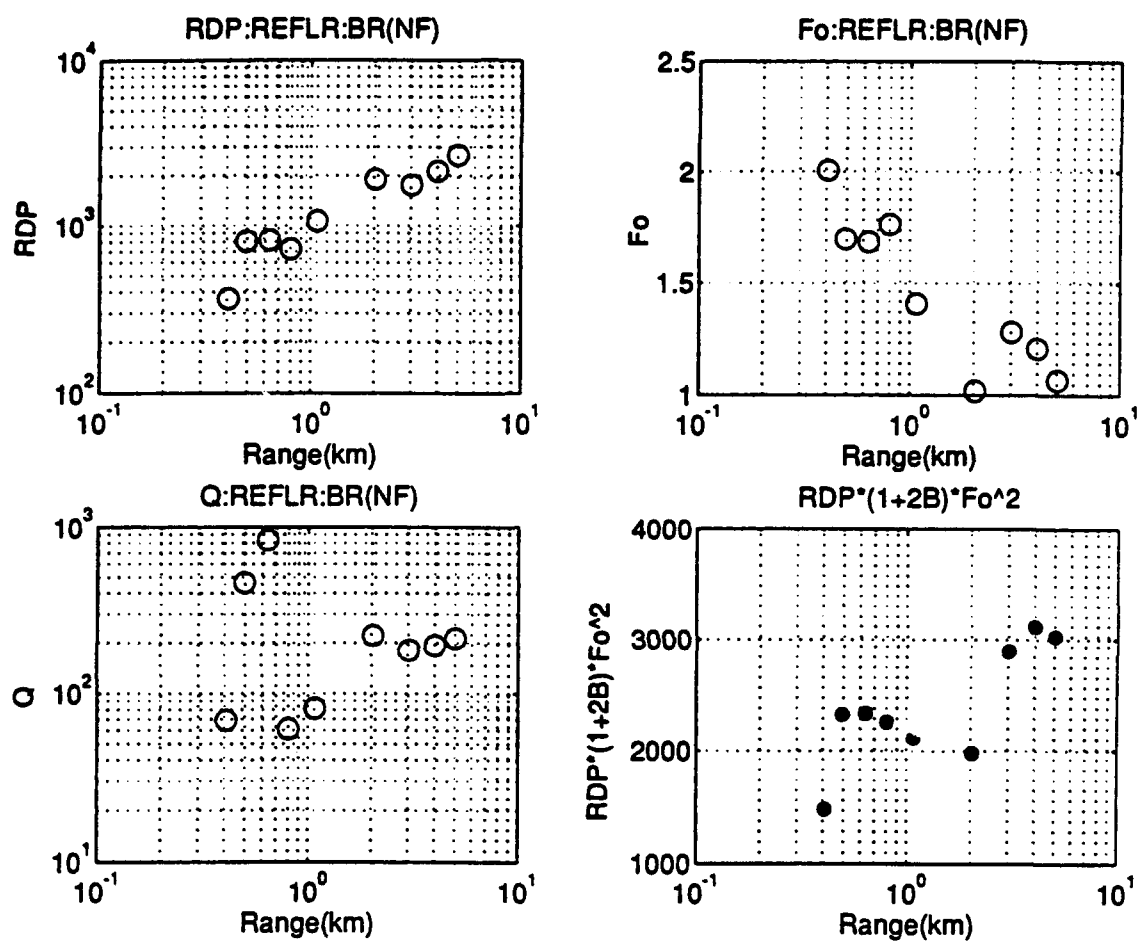


Figure 3.50 Parameter estimation with Brune's model. Homogeneous full-space path correction was made. Radial component is used.

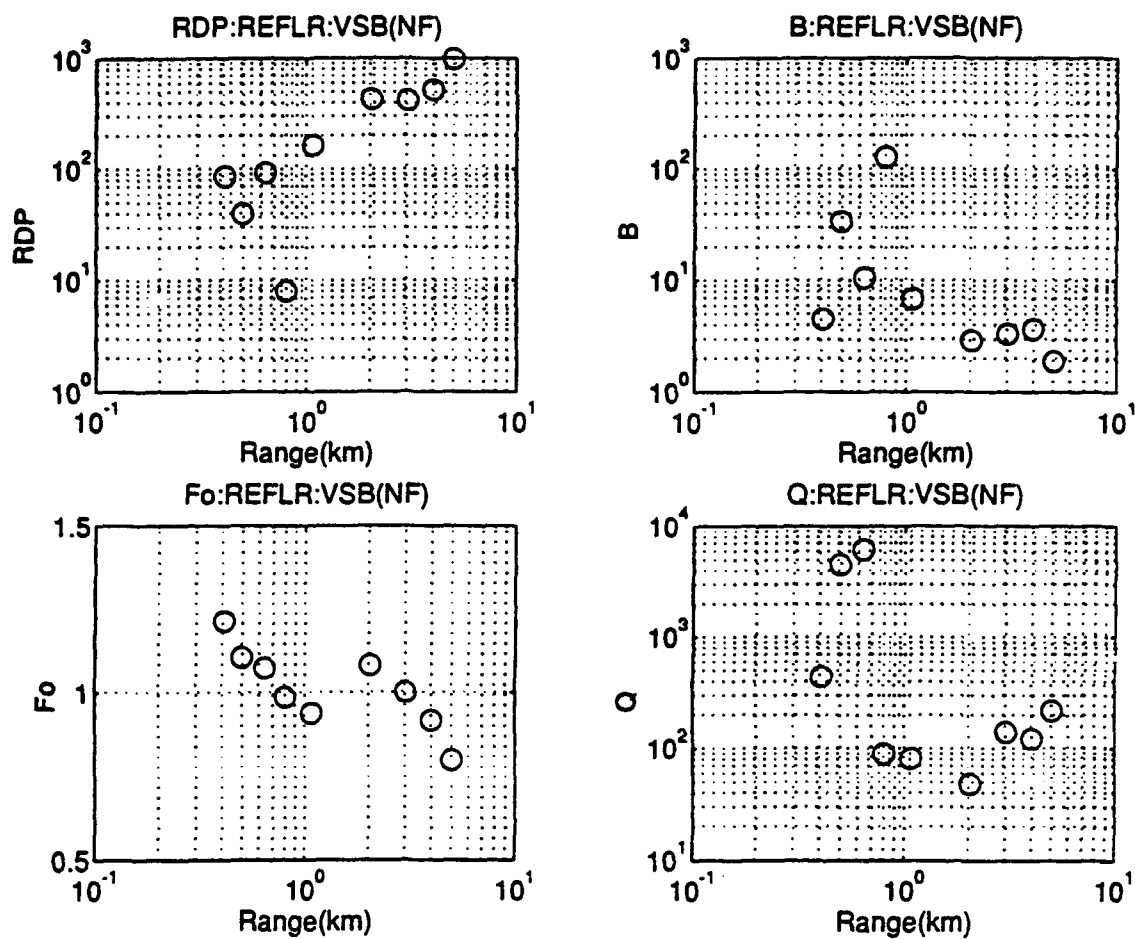


Figure 3.51 Parameter estimation with von Seggern-Blanford's model. Homogeneous full-space path correction was made. Radial component is used.

interaction was not considered. The increasing trend of steady state RDP with decreasing corner frequency seems to be the effect of surface waves that have a smaller spreading rate than the body waves used in the forward model. In the VSB model inversions, overshoot shows very high values (several hundreds) due to the surface wave effect. There is a trend that overshoot decreases as distance increases beyond 1 km, but it is not clear that this trend is consistent after 5 km. Excessive surface wave contribution, whose effect is dominant beyond the corner frequency, cannot be explained by the P-wave forward model. In this case, the value of estimated overshoot is not a true overshoot, but the compensation for the unmodelled surface wave effect.

Inversion was performed to check the applicability of proper path correction. Homogeneous half-space path model was used as a path correction. The physical properties at the source depth ($V_p=2.2$ km/sec, density= 1.85 gram/cm³) were assumed as an average value in a homogeneous half-space. Figure 3.52 and 3.53 show the results of inversion when Brune's model and von Seggern-Blanford's model were used as a forward model. Only the radial components were analyzed since these are the least contaminated by the surface waves. When the Brune's model was used as a forward model (Figure 3.52), steady state RDP decreases as the distance increases within 1 km. This is the region where the direct body wave is dominant. The estimated corner frequencies are greater

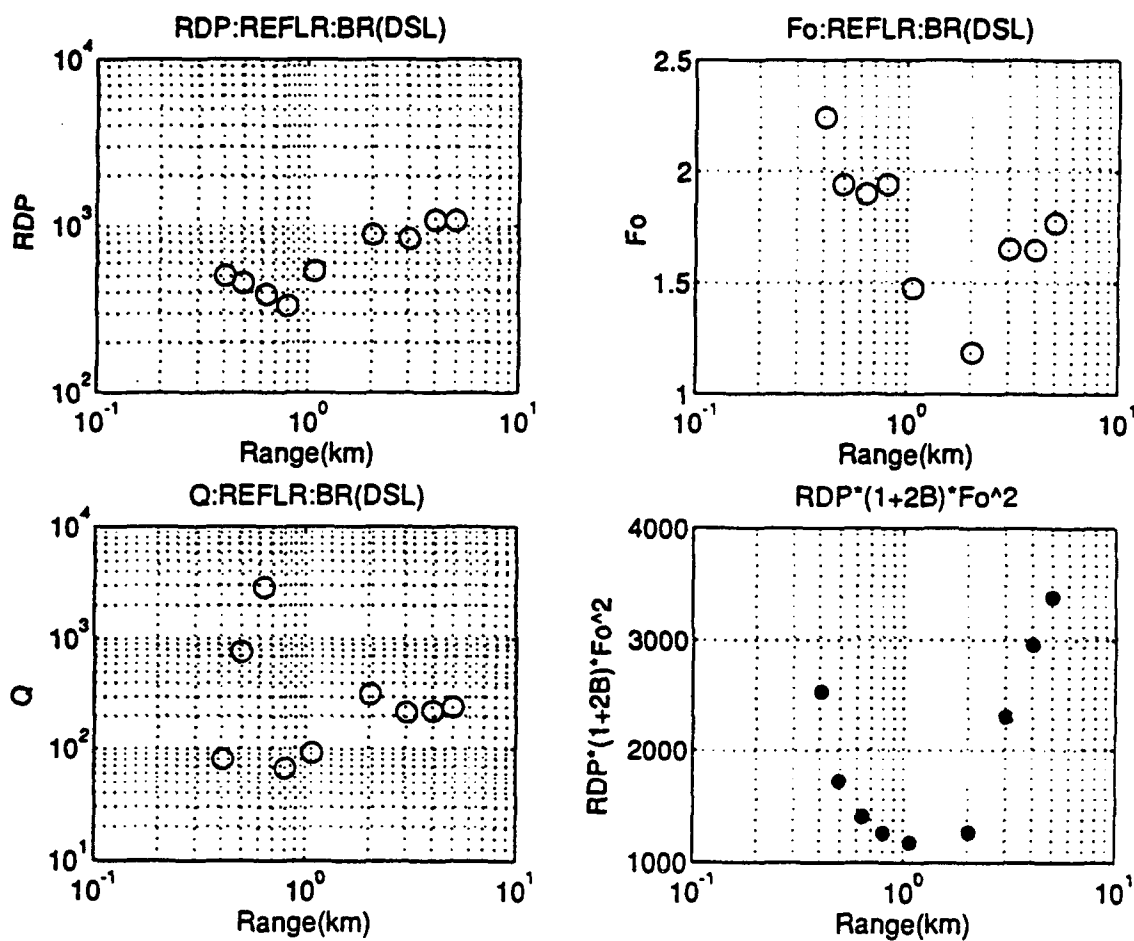


Figure 3.52 Parameter estimation with Brune's model. Homogeneous half-space path correction was made. Radial component is used.

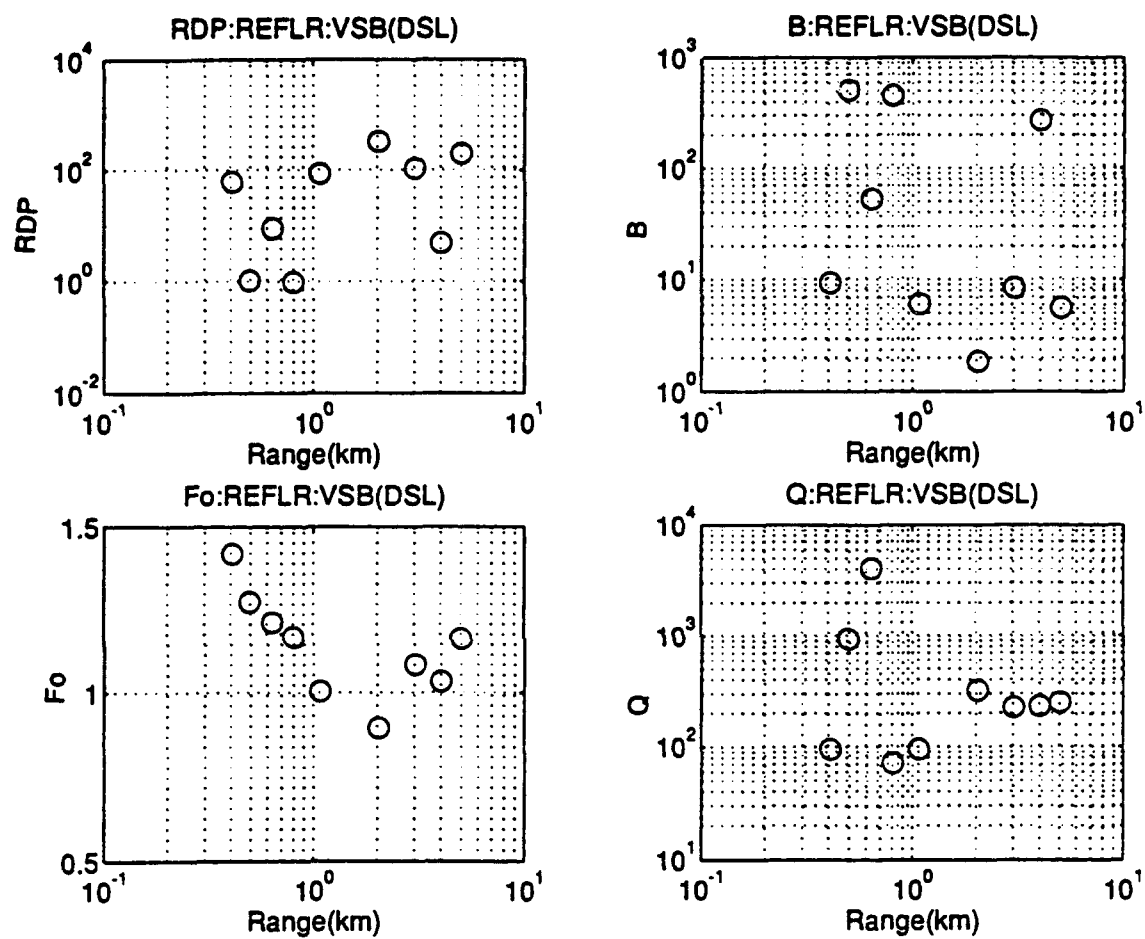


Figure 3.53 Parameter estimation with von Seggern-Blanford's model. Homogeneous half-space path correction was made. Radial component is used.

than the applied corner frequency(1.5 Hz) in this region since Brune's model was used as a forward model while the data include the overshoot effect. It varies from 1.5 to 2.3 Hz in this range. Estimation of Q is almost meaningless since both the synthetic data and the correction term(homogeneous half-space path) have large values of Q . Ideally, estimated Q should be infinity which is impossible in the numerical analysis for convergence. Therefore, there is always a possibility of insufficient convergence in parameter estimation in this data set, otherwise the inversion either diverges or iterates towards a Q of infinity. The high frequency approximation shows that the combined effect of steady state RDP and corner frequency tends to decrease due to the free-surface interaction first and increase due to the surface wave effect. The surface wave effects(both amplitudes and dispersion) cannot be compensated thoroughly by the homogeneous half-space correction.

When the von Seggern-Blanford's forward model was used(Figure 3.53), the results show a large variance in parameter estimation. The main reason of large variation in steady state RDP and overshoot is in the numerical noise added to low frequencies. Large fluctuation of overshoot and corresponding STEADY STATE RDP in Figure 3.53 shows that the source models with overshoot(VSB, HH, and HS models) behave ill-posedly around the corner frequency.

One of the main purpose of synthetic test is to check the applicability of ω^{-3} model for data from ω^{-2} model. When Helmberger-Hadley model was used as a forward model, inversion diverged in most cases owing to the slower decay rate in the high frequency data which cannot be expressed with ω^{-3} model and real Q .

Data with large attenuation

The same reflectivity synthetic seismograms were generated with large attenuation ($Q_\alpha=30/Q_\beta=17$). Figure 3.54(a and b) shows the synthetic velocities. Each seismogram is normalized by its own maximum amplitude. Epicentral distances range from 0.1 to 20 km. To reduce the wrap around phase before the first arrival, the time duration was doubled. However, there still exist the truncated phase at large distances. Brune's and Helmberger-Hadley's model without overshoot were used as a forward model. As a path correction, homogeneous-half space and layer over half-space models were used for the Brune's model application and layer over half-space model was used for the Helmberger-Hadley model application in inversion. Homogeneous full-space path model was not used because of its inapplicability as shown in the previous section.

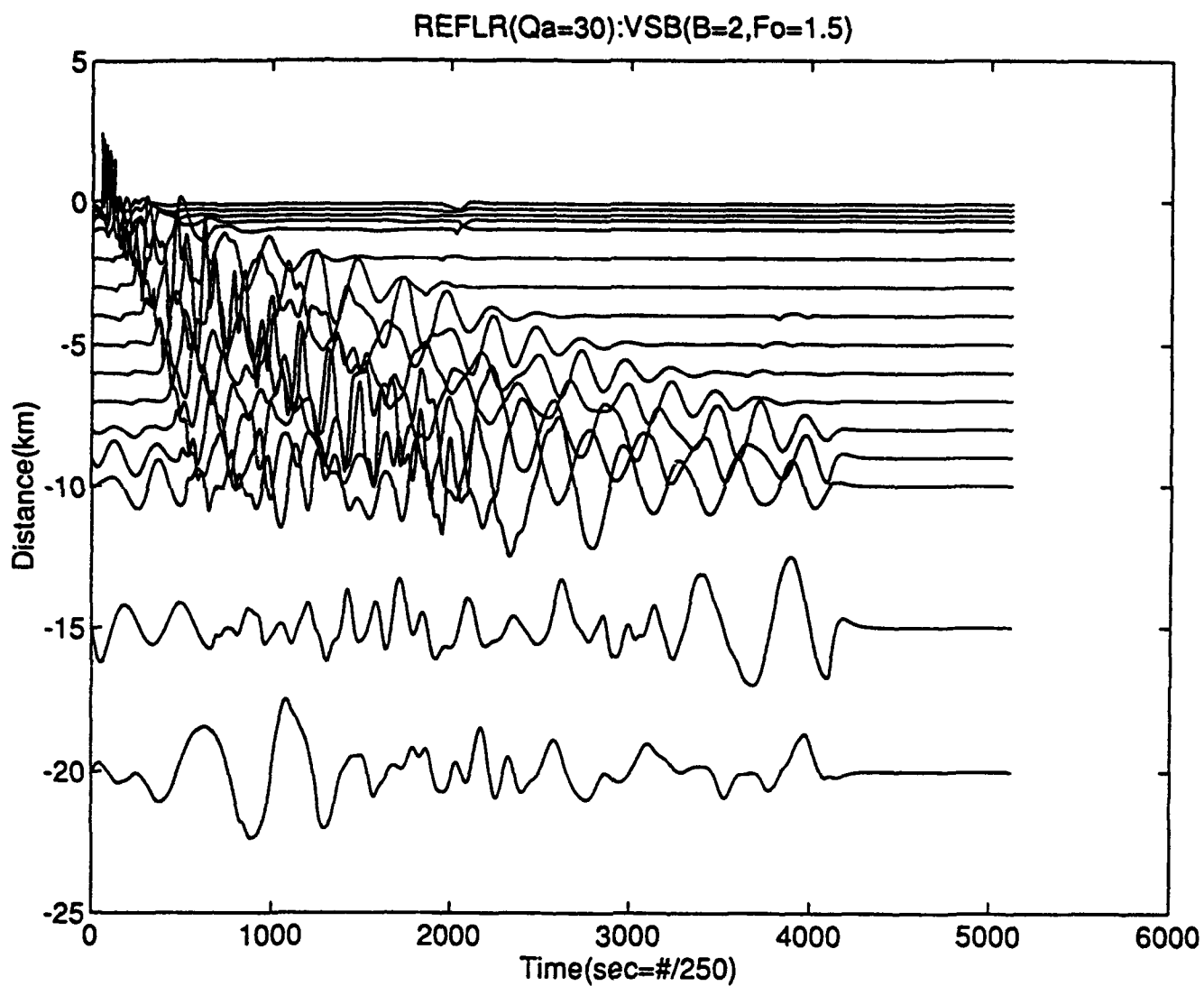
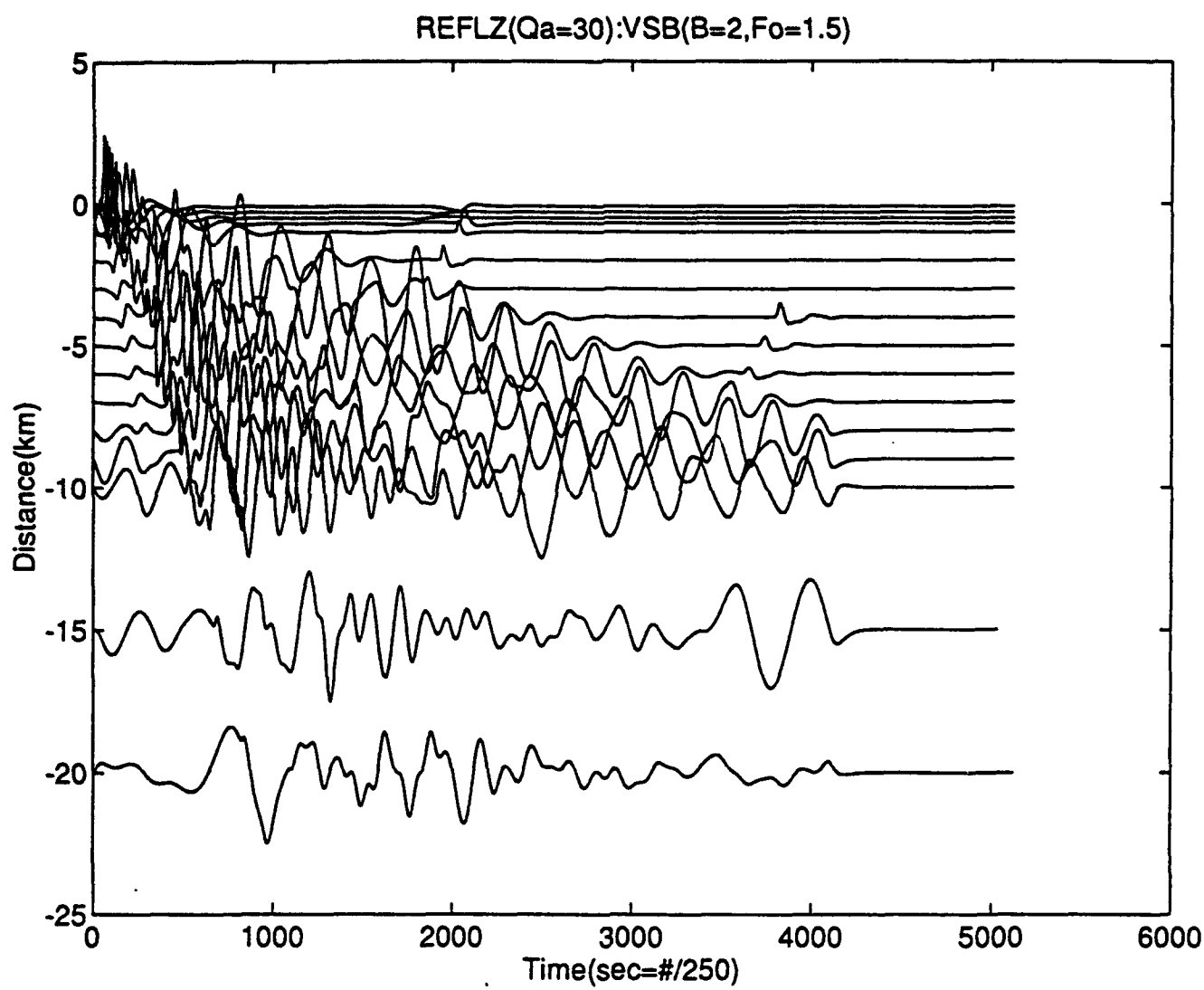


Figure 3.54 Synthetic seismograms with large attenuation($Q_a=30$). The VSB model($B=2, F_0=1.5$ Hz) were convolved. (a)radial; (b)vertical.



(b)

Brune's model

Figure 3.55 shows the result of inversion of the radial component with the homogeneous half-space path correction(model 1). Both data and path model were normalized to the response of the same moment (10^{20} dyne-cm). steady state RDP within 1 km shows the same trend as the previous analysis with small attenuation. Expected steady state RDP is 1.7 m^3 and expected corner frequency is 2.8 Hz from Figure 3.24. Steady state RDP shows almost constant values between 1 to 8 km with the value of around 8 m^3 and drops beyond that distance as the distance increases. Corner frequency is almost constant(mean=1) with small variation which increases at large distances. Compared with the result with small attenuation(Figure 3.52), corner frequency estimation shows quite a different value. It seems to be the result of an insufficient convergence in case of data with small attenuation. Estimated corner frequency with large attenuation shows even smaller value than the applied corner frequency with overshoot while the Brune's model was used as a forward model. This implies that the estimated corner frequency is biased strongly by the surface waves. The biases of the overshoot and corner frequency affect the estimation of steady state RDPs also and resulted in an overestimate of steady state RDP. Q generally increases as the distance increases beyond 1 km. Within 1 km of epicentral distance where the surface wave generation is not

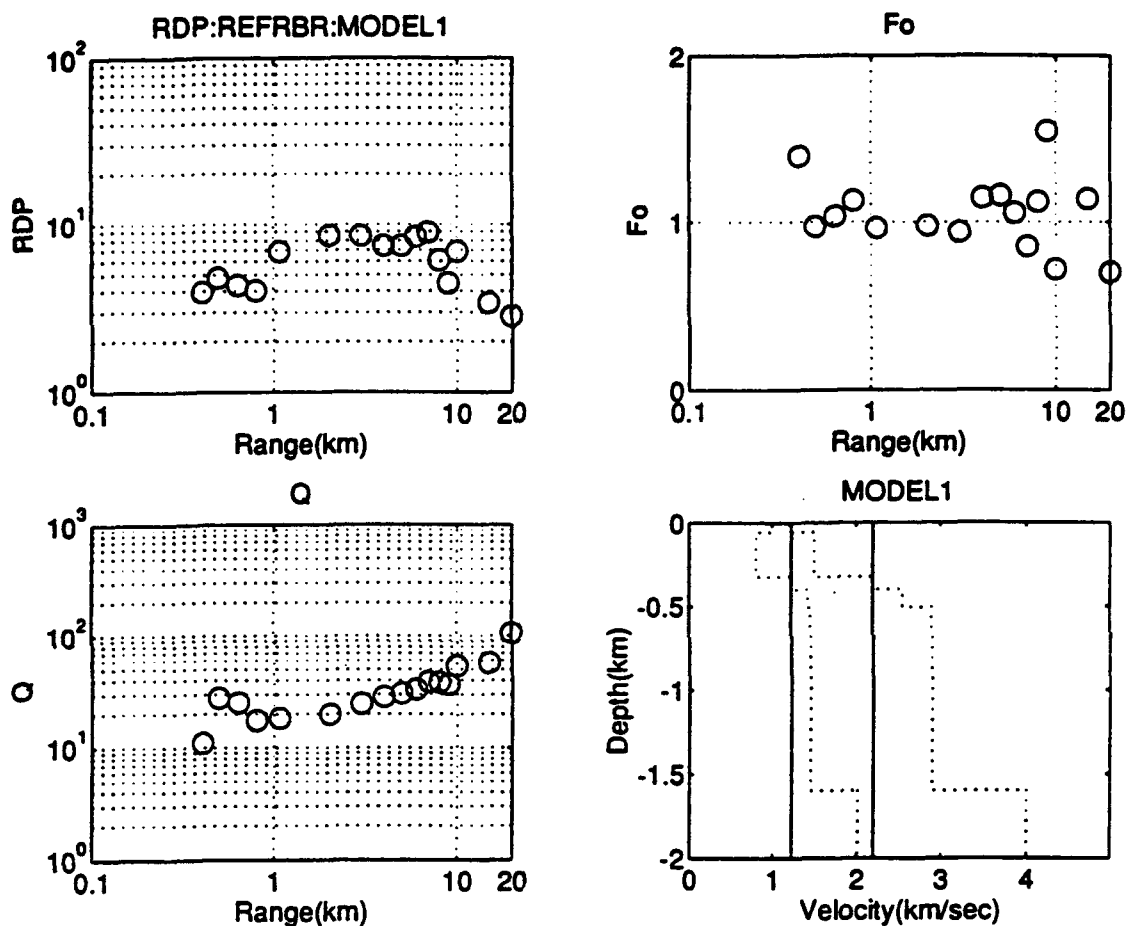


Figure 3.55 Parameter estimation from the radial component with large attenuation. The Brune's model with homogeneous half-space path correction was used.

so significant, the free surface interaction does an important role to determine parameters. While the incident angles of the synthetic data do not change much, the incident angles of path correction change a lot at this range, which leads to inconsistent path correction. At large distances, the data are seriously contaminated by the surface waves while the path correction maintains relatively stable body waves. In the model plot in Figure 3.55, the solid lines are the P- and S-wave velocity model used for the path correction while the dotted lines are the velocity structure used to derive the synthetic seismograms. The effect of inconsistent correction for the path effect is clearer in the analysis of vertical component(Figure 3.56). Relative amplitude of the surface waves compared to body waves is much higher in case of homogeneous half-space than in the data from the layered structure because almost all the body waves travel radially at large distances in the homogeneous half-space. The inconsistency of the path correction results in an unreliable estimation of source parameters.

When the layer over half-space model(model 2) was used as a path correction, the estimation shows consistent result in its steady state RDP in case of radial component(Figure 3.57). Eleven point running average was applied to the path model to make the spectra smooth. Corner frequency is almost constant(except the case of $R=0.1$ km) in the range of less than 8 km and drops consistently beyond that as the distance

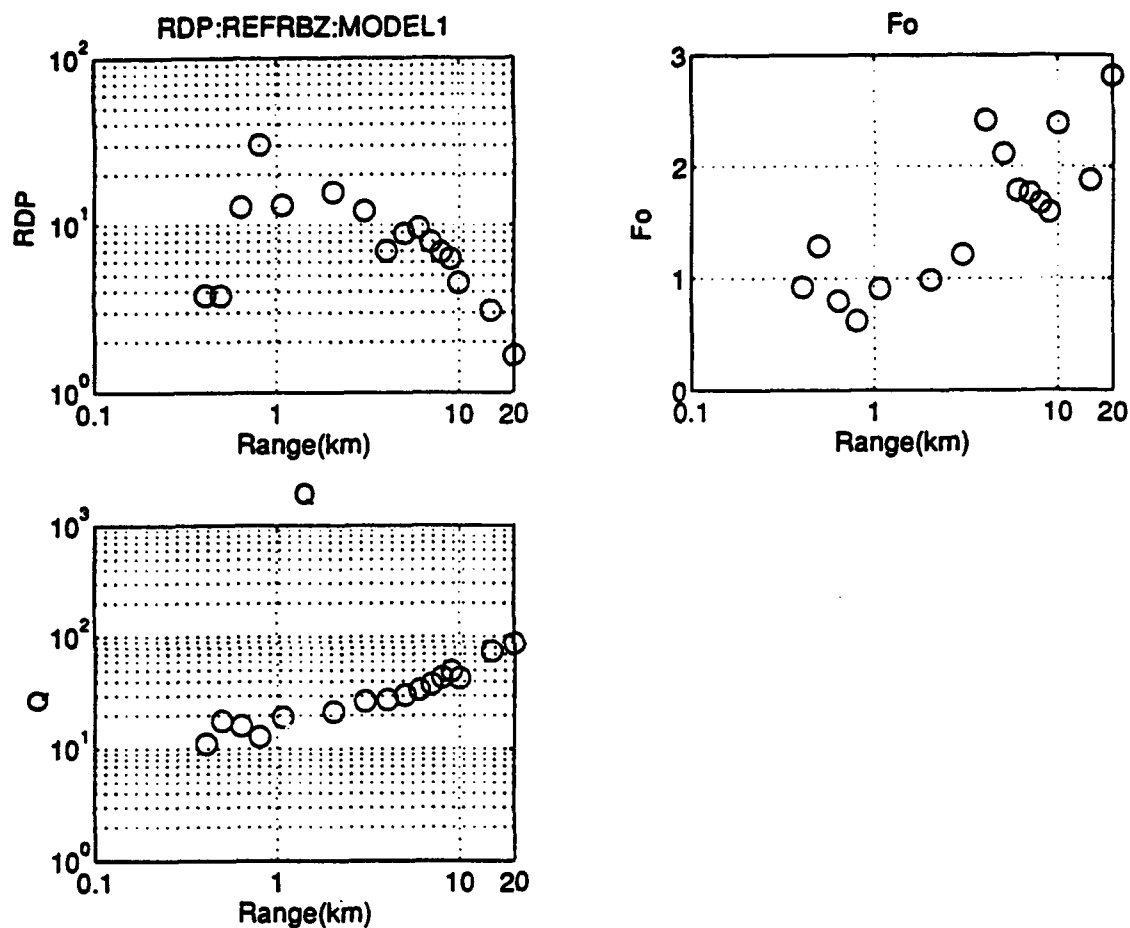


Figure 3.56 Parameter estimation from the vertical component with large attenuation. The Brune's model with homogeneous half-space path correction was used.

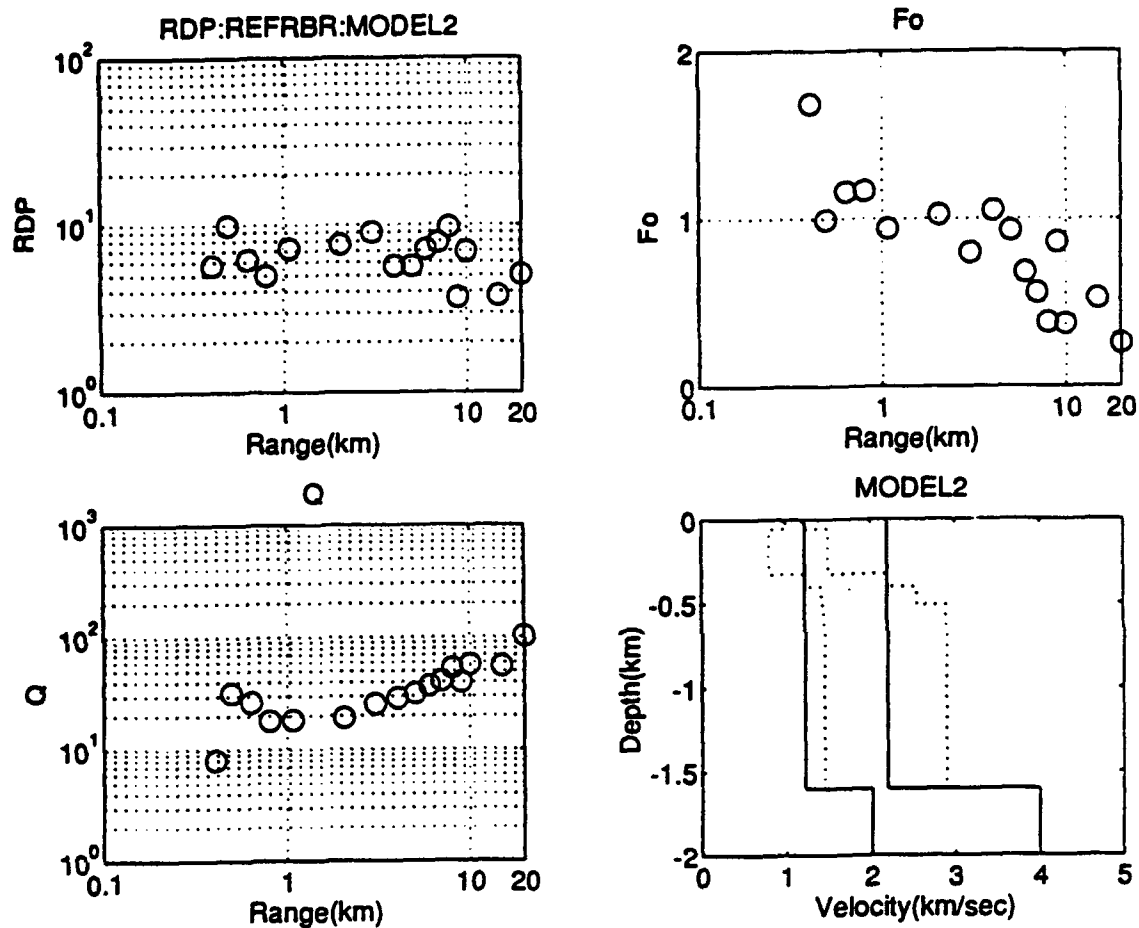


Figure 3.57 Parameter estimation from the radial component with large attenuation. The Brune's model with layer over half-space path correction was used.

increases since the surface wave effect becomes more and more dominant in the spectra. At large distances beyond 8 km, the corner frequency represent the corner frequency of surface waves rather than that of body waves. The analysis of vertical seismogram shows an even worse estimate than the radial component(Figure 3.58) since the vertical component is contaminated more by the surface waves. The free surface interaction, on the contrary to the homogeneous half-space correction, is not a major factor on the bias of source parameters since the layer over half-space has almost the same incident angle at the free surface and shows constant steady state RDP estimates. Q increases consistently as the distance increases beyond 2 km. The increase of Q with respect to distance is related to the difference between the actual travel time and the assumed travel time. As the distance increases, the waves travel through the higher velocity layer while the assumed travel time is based on the constant velocity at the source region. Assumption of larger travel time is compensated by the larger Q since the real estimated value in an inversion for the frequency independent Q model is the value of Q divided by travel time. P-wave velocity was assumed as 2.2 km/sec from the surface to the depth of 1.6 km and 4.0 km/sec below that and the corresponding S-wave velocity and density were used for the generation of the Green's function in the layer over half-space.

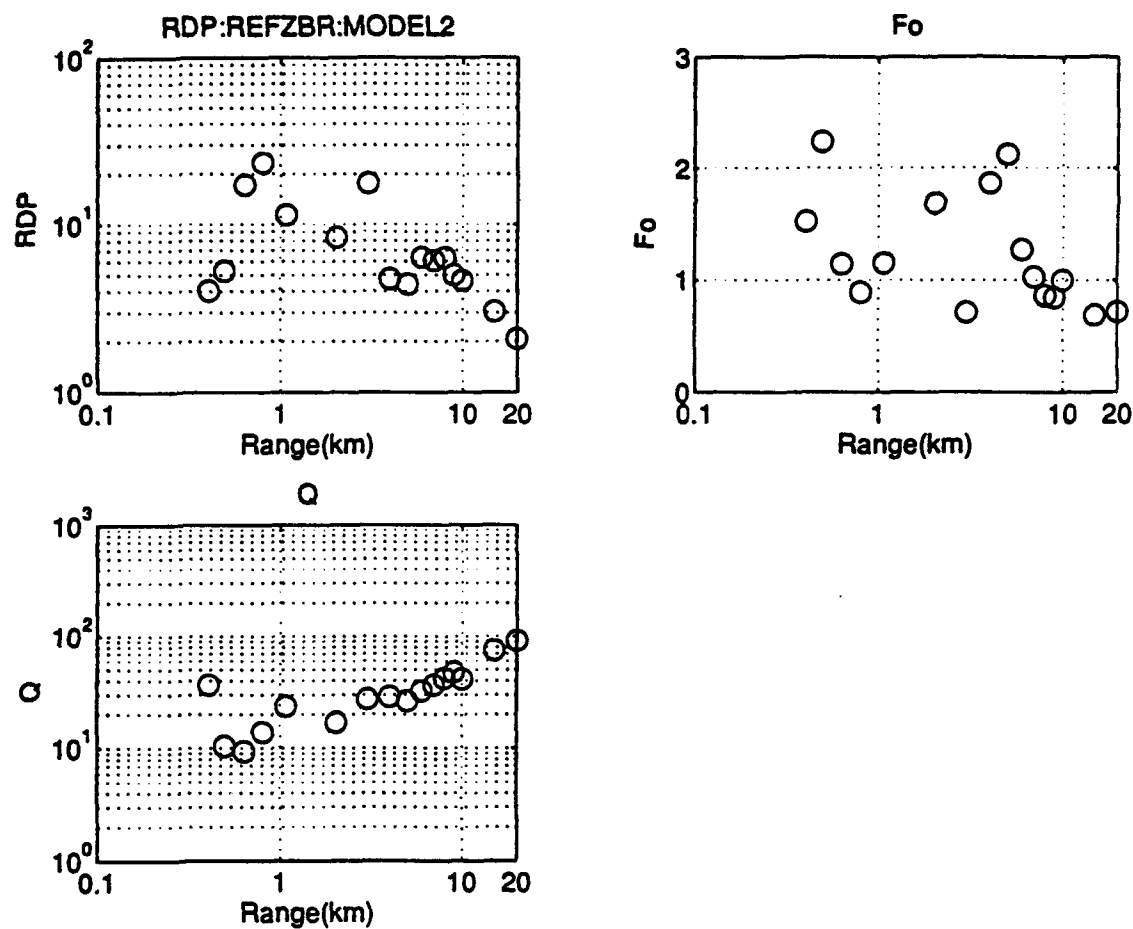


Figure 3.58 Parameter estimation from the vertical component with large attenuation. The Brune's model with layer over half-space path correction was used.

It is impossible to resolve source parameters accurately without knowledge of the underlying structure. Needless to say, the more accurate path corrections are indispensable to estimate source parameters. From the above tests with various types of path correction, the assumption of homogeneity of the space is not sufficient to resolve source parameters from the simple path effect.

Helmberger-Hadley model

The same data were inverted with the Helmberger-Hadley model without overshoot to check the trade-offs between the source model and Q in the near-source region. The layer over half-space path corrections were used for the data analysis.

Figure 3.59 shows the result of inversion with both forward models. Radial components were analyzed. The thick lines are the result of inversion with the Helmberger-Hadley's model without overshoot and the thin line is the result of Brune's model. Data are also plotted as a thin solid line. As shown in this figure, both forward models (Brune's and Helmberger-Hadley's model without overshoot) inverted successfully. Figure 3.60 shows the estimated parameters by both models. When Helmberger-Hadley's model was used as a forward model, Q shows large values at short distances. The effect of real $Q(Q_\alpha=30)$ was traded off with the high frequency roll-off in ω^{-3} model and estimated Q shows large value. However, as the distance

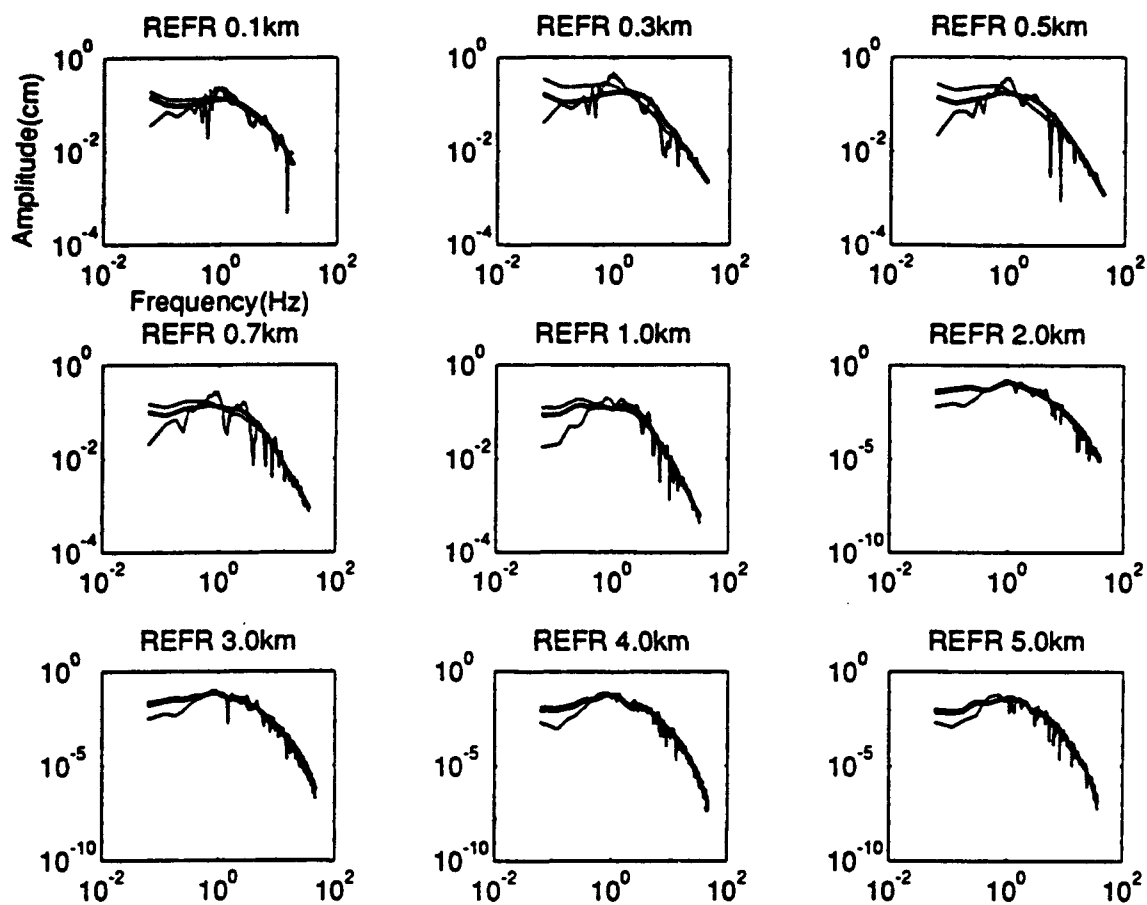


Figure 3.59 Result of inversion with forward models. Thick smooth lines are the result of inversion with the HH model and the thin smooth lines are the result of inversion with the BR model. Data are also plotted. The difference between the result of two models becomes obscure as the distance increases.

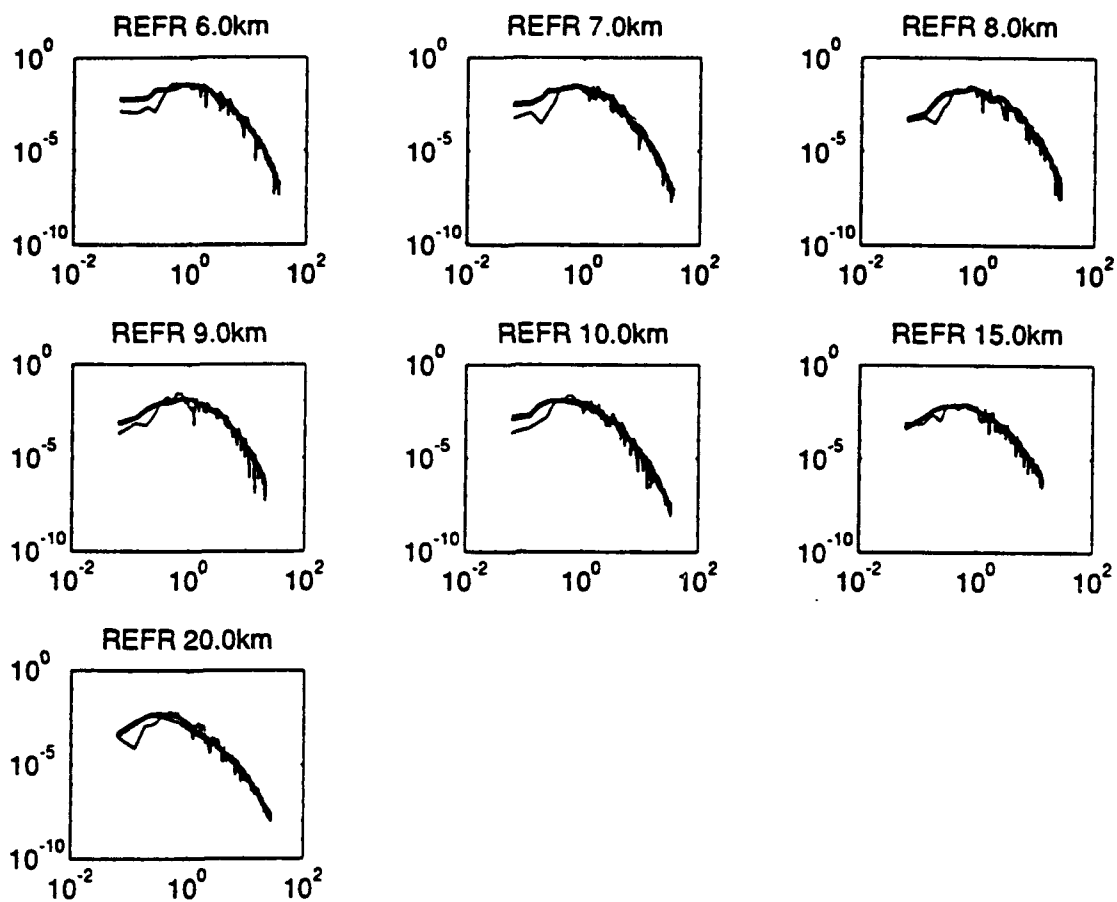


Figure 3.59 Continued.

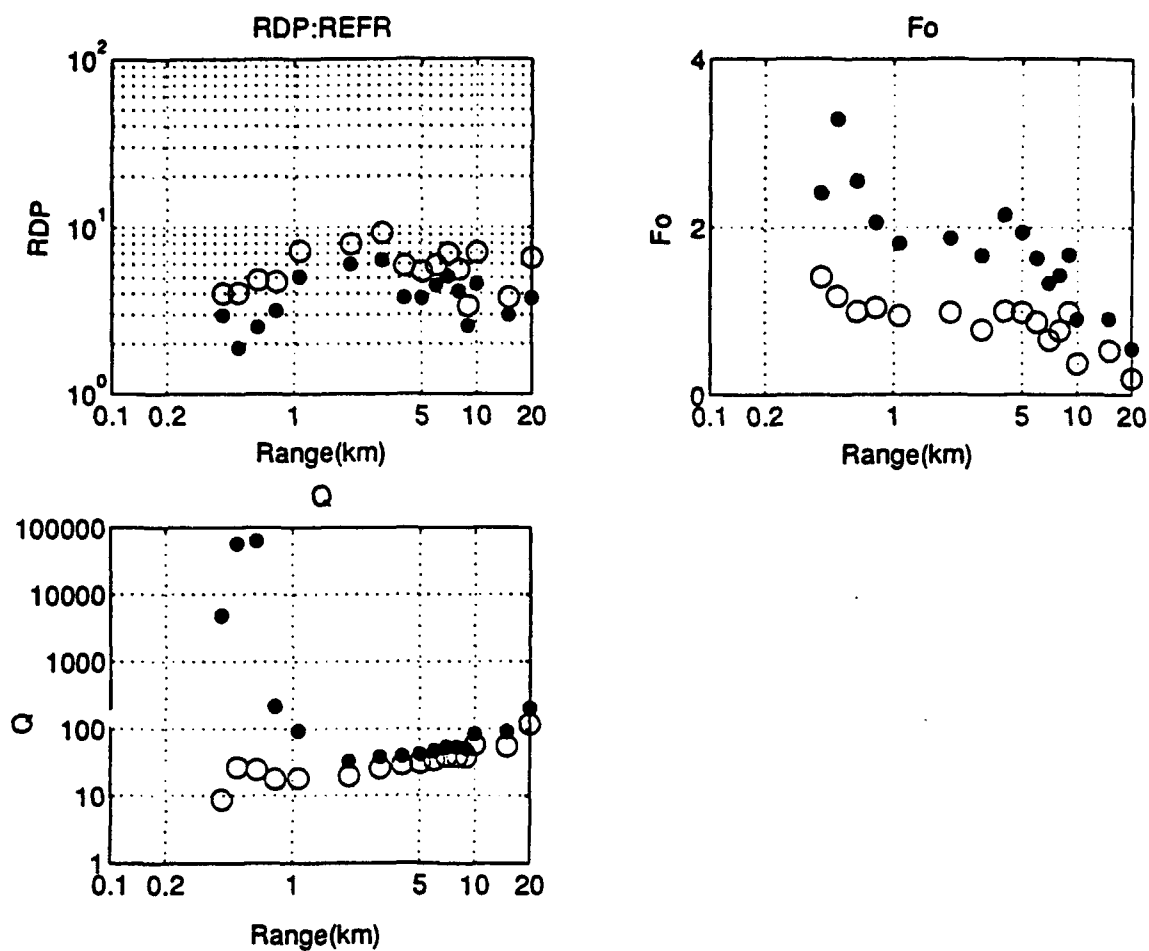


Figure 3.60 Plot of estimated parameters by both models(HH and BR). The HH model generally shows largr corner frequencies. The HH model shows unrealistic high Q at very short range.

increases, the development of surface wave biases the shape of the spectra, which is sufficient to represent the data by the ω^{-3} model. At this range the estimated Q is close to the expected value. The standard deviations of the applied models (Figure 3.61) also indicates that it is almost impossible to identify the source model when the attenuation is large. In Figure 3.61, the standard deviations by the Brune's model were expressed as positive while those by the Helmberger-Hadley's model were expressed as negative for the comparison. The changes of the standard deviations from receiver to receiver are mainly attributed to the number of data points used in an inversion. Even though both models represent the data well and there is no preference between the models in the statistical sense, there is a big discrepancy in the Q estimation at short ranges when wrong model was applied. Without considering path effect such as surface waves, it is impossible to select the proper explosion source model.

There is no known way to separate the body and surface waves where the dispersion of the surface waves is not developed fully. It is not possible to separate body waves and the surface waves in the time domain by using techniques such as polarization analysis in the near source region since the body and surface waves interfere. In an area where the attenuation is large it is especially important to obtain very near-source data to select a proper source model.

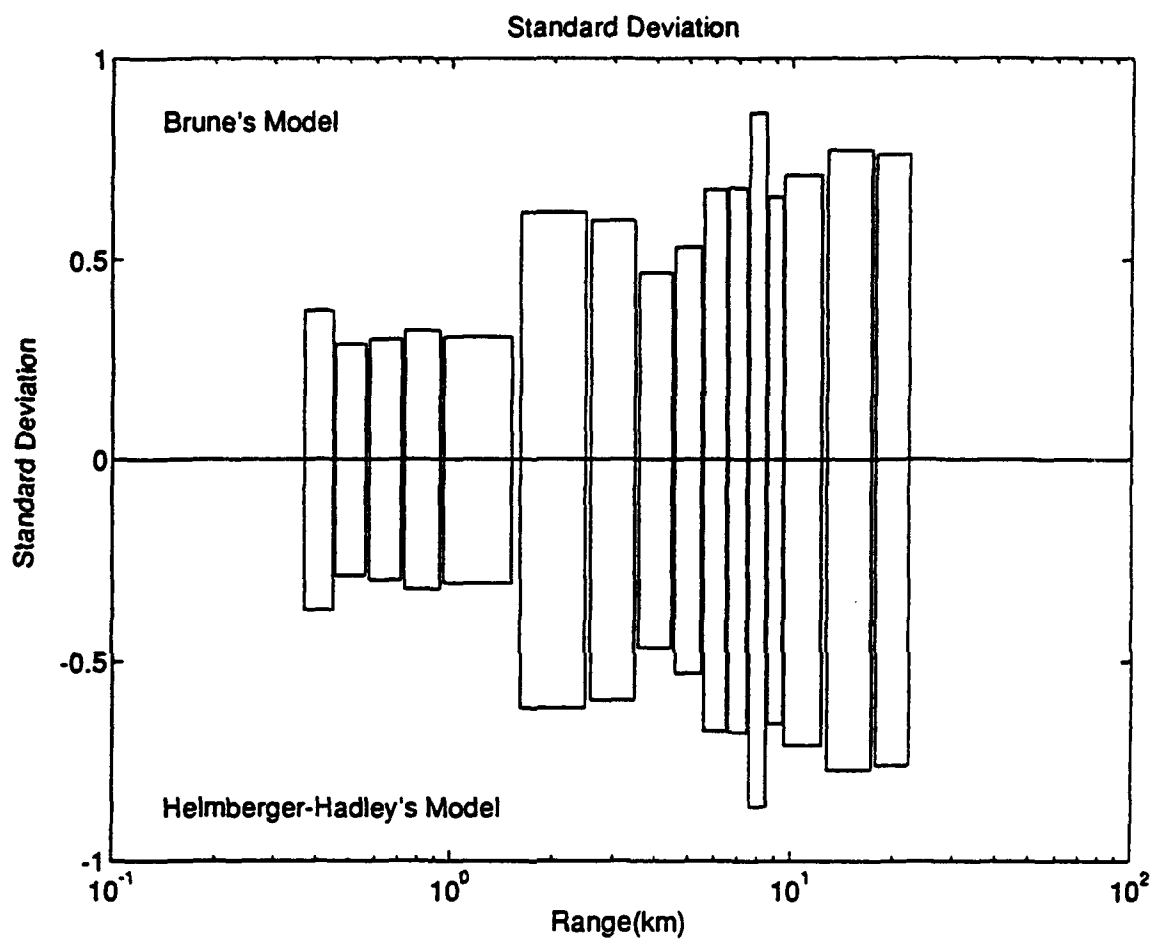


Figure 3.61 Standard error, standard deviation of the logarithm of the residuals, by the VSB and the HH model cannot be a criteria to characterize the source function. Standard deviations by the VSB model are plotted in the positive side while those by the HH model are plotted in the negative side.

Bootstrap Statistics with Data Including Outliers

Since it was proven in the statistical field (Efron and Tibshirani, 1986) that the bootstrap method is powerful to analyze the limited data set which is the case of the estimation of steady state RDP, overshoot and corner frequencies in source parameter inversion, the only thing left is to validate the behavior of the bootstrap with the data including a few outliers. As was introduced and mentioned in the previous chapter, bootstrap with Monte Carlo method can be combined with the non-linear least-square inversion. It may be called non-linear algorithm with bootstrapping.

Even though it is generally accepted that the bootstrap is a powerful tool in the statistics, the applicability is not fully investigated yet since it is relatively new technique. After the bootstrap was introduced by Efron in 1977, there were some applications in seismology for the estimation of seismic properties, but their applications were limited to the basic properties of bootstrap. What we found in the application of bootstrap is that the bootstrap estimates are also useful in estimating parameters from the data including some outliers. In testing the applicability of the bootstrap method to the data with outliers, empirical examination was performed with the synthetic data generated by the computer.

One-Dimensional Statistics

For one-dimensional bootstrap test, 100 normally distributed random numbers were generated (Figure 3.62). The histogram shows that the distribution of the sample shapes Gaussian. The mean and the standard deviation are zero and 0.9928 respectively. Mean was used as a parameter for a test. The process of bootstrap test is, as was explained in the previous chapter, i) sample 100 times from the data with the concept of Monte Carlo method with replacement, ii) take the mean from the bootstrap samples, and iii) repeat the same process several times. Then we get new data set which is composed of mean values of each bootstrap samples. Figure 3.63 shows the process of bootstrap estimate. Bottom figure in Figure 3.62 shows the distribution of the mean of the bootstrap samples. The mean of the data is presented as an open circle while the mean of the bootstrap mean is presented as a cross. As a reference median of the data and the mean of the bootstrap are denoted as a cross and a plus sign respectively. Cumulative distribution function down in the figure also shows the symmetric shape. An asymmetric ratio, which is defined here as

$$\text{Asym. Ratio} = \frac{(95 \text{ percentile} - 50 \text{ percentile})}{(50 \text{ percentile} - 5 \text{ percentile})}$$

is 0.9 which is close to one for the perfect symmetry. The distribution of the bootstrap mean is also zero like that of

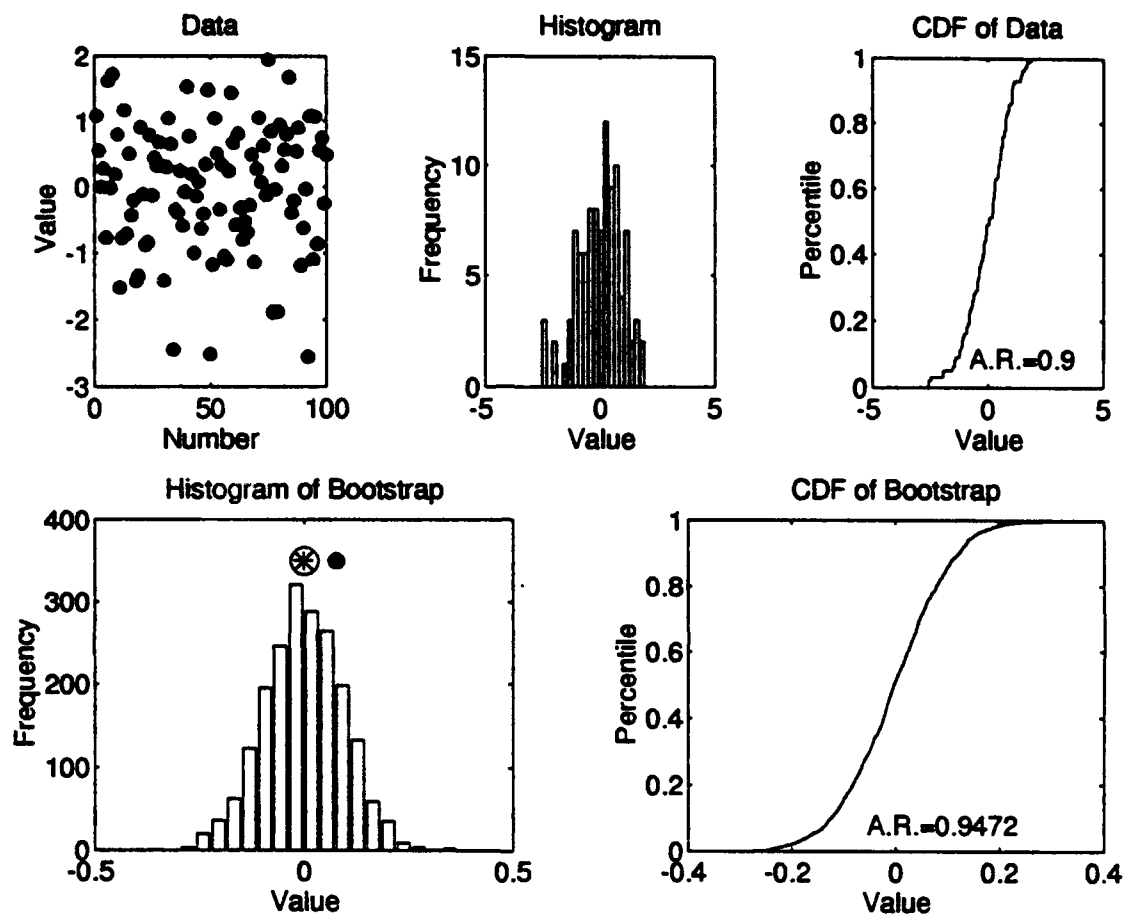


Figure 3.62

BOOTSTRAP PROCEDURE

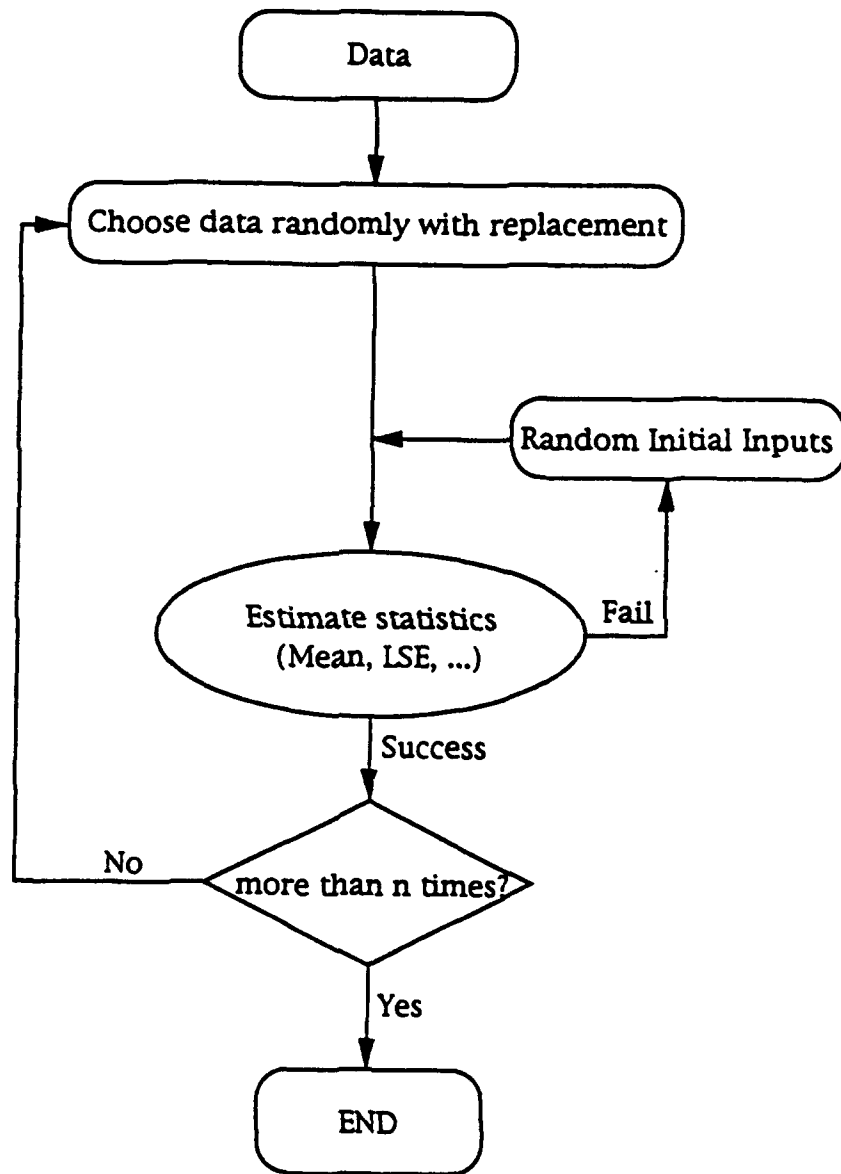


Figure 3.63 Bootstrap procedure

the data. Actually, there is no reason to apply bootstrap method in this case since the distribution of the population is known as Gaussian and the data itself are designed to represent the distribution of the population well.

To validate the reliable estimate of the statistics from the biased limited samples, three outliers are inserted to the data(Figure 3.64). The mean of the data with outliers is still the same as those of the original data since the mean of the outliers is adjusted to zero. Figure 3.64 shows the result of 2000 bootstrap calculations. The distribution maintains the normal shape of the distribution and any other statistics used here do not change much from those of the original data. If one of the values of outliers is changed so that the mean of the data is affected by these outliers, the distribution of the bootstrap mean is not Gaussian but shows skewed distribution(Figure 3.65). Naturally, an asymmetric ratio also moves away from the symmetry. The direction of the tail of the bootstrap distribution is toward the maximum value of outliers. While the trend of skewness increases as the effect of outliers increases, the distribution of the bootstrap starts to show multiple peaks after a certain limit(Figure 3.66 and 3.67). Observations of the empirical test display there is a periodicity in the peaks. Two of the peaks in the distribution of the bootstrap are corresponding to the median and the mean of the data as shown in the figures as a dot and an open circle. The tail

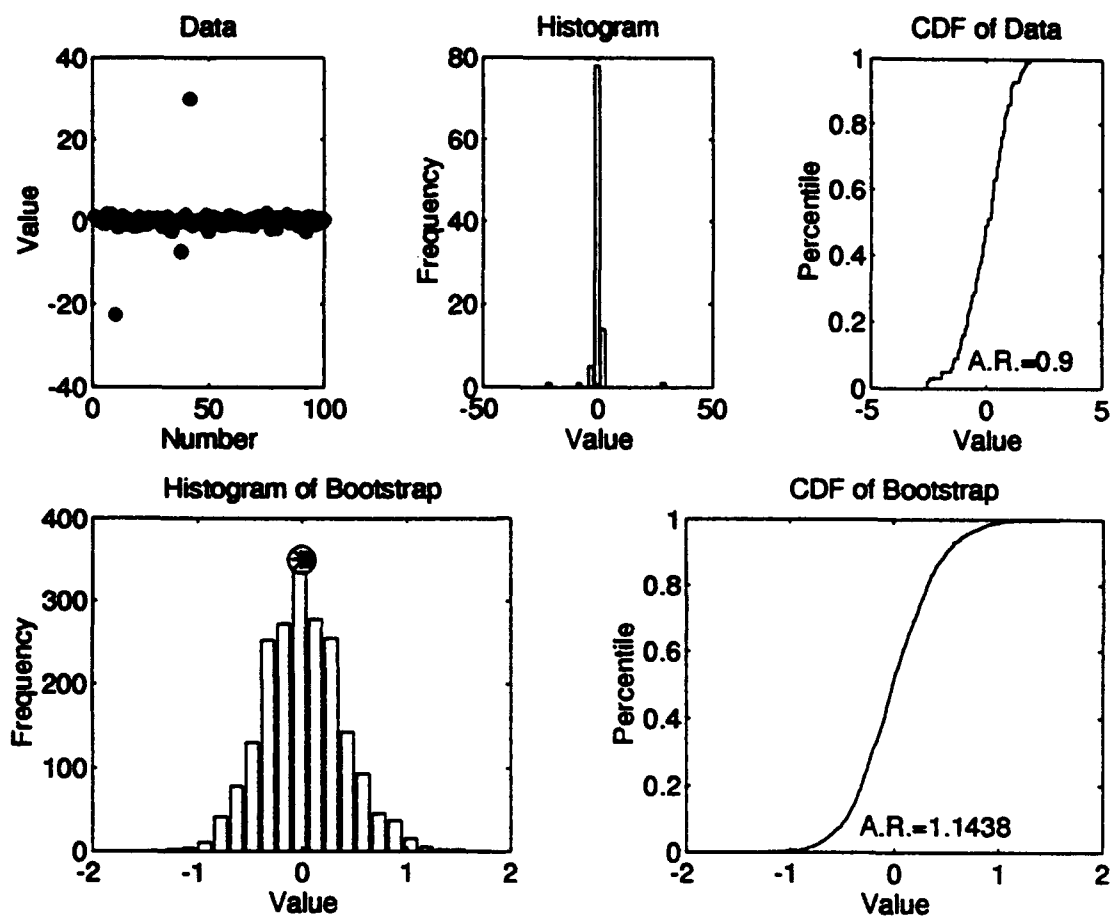


Figure 3.64

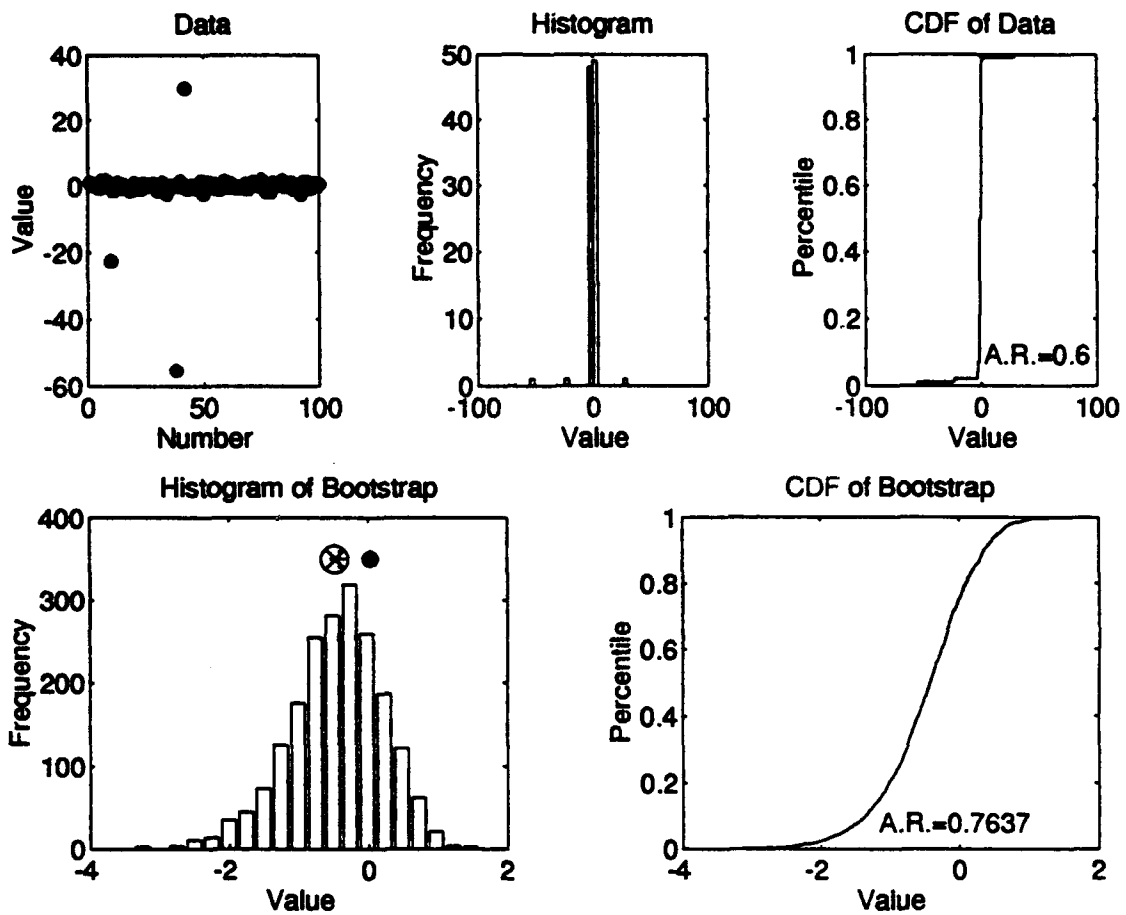


Figure 3.65

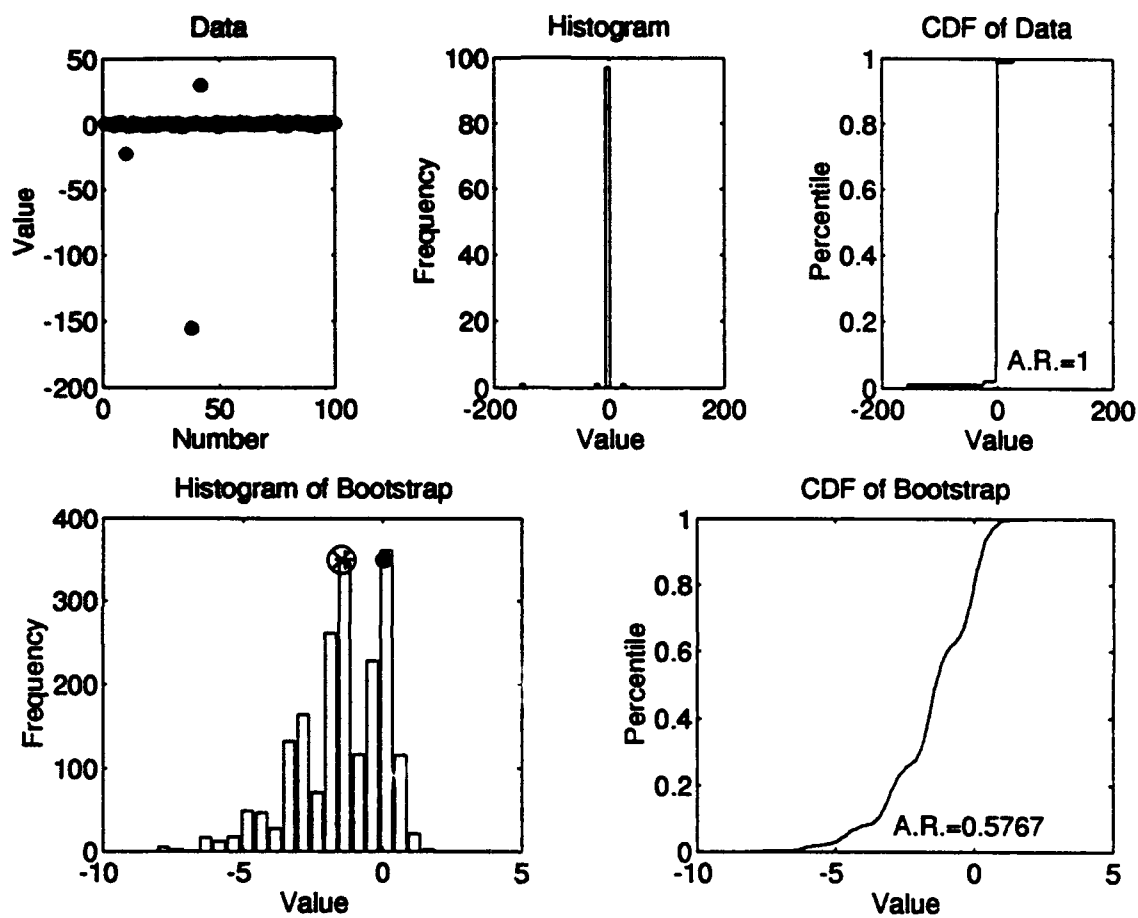


Figure 3.66

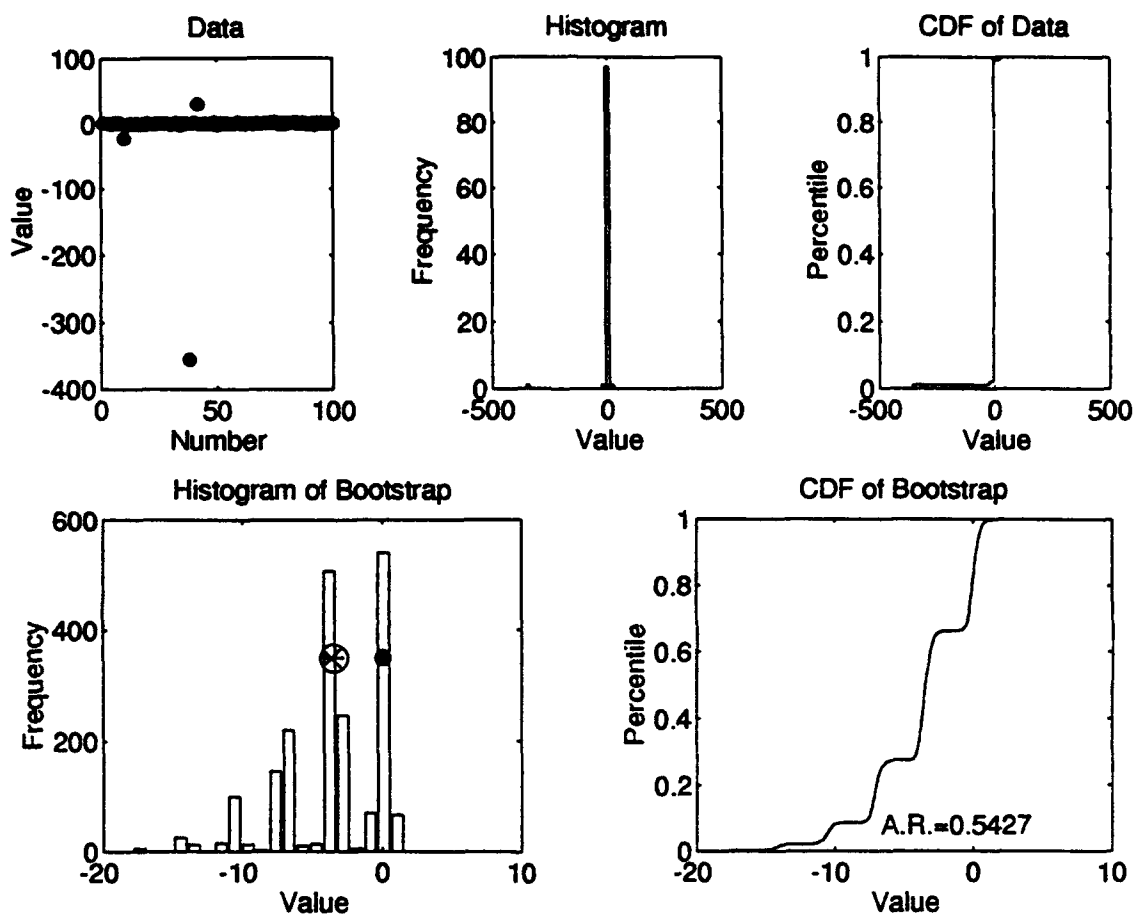


Figure 3.67

of the envelope of whole distribution is elongated to the direction of extremal size of outlier.

This distinctive characteristic of the bootstrap with big outliers can be explained as follows.

1) Bootstrap distribution is finite discrete. Its expected minimum and maximum values of bootstrap mean are also finite.

$$\lim_{B \rightarrow \infty} \{\min(E_B[X])\} = \min(\text{data})$$

$$\lim_{B \rightarrow \infty} \{\max(E_B[X])\} = \max(\text{data})$$

where B is the number of bootstrap process.

2) There is a probability that the bootstrap sample does not take any specific data points. This probability has its own distribution. Therefore, there is a distribution which does not take any outliers.

3) The mean of the bootstrap sample shifts mainly by the outliers.

Since the distribution of the bootstrap mean is the summation of individual distribution whose mean is mainly controlled by the outliers and since the distribution of bootstrap is finite, there exist finite number of peaks which represent the mean of the bootstrap samples excluding specific data points. If the outliers lie far beyond the standard deviation of the rest of the data so that the bias of the mean is large fraction of the standard deviation of the data without outliers (in this test, more than 20 %), the

interference between individual distribution is small and shows peaks. When the values of the outliers are not so distinctive that they cannot bias the mean no more than a fraction of the standard deviation of the data (in this empirical test, bias of the mean within 20 percent of the standard deviation), individual peaks cannot be seen due to the interference of each distribution. When the number of outliers is increased, the probability to sample outlier-free bootstrap data is slim. In this case, the distribution of the bootstrap mean generally does not show individual peaks but skewed distribution since the shift of the mean is small.

One-dimensional bootstrap analysis can be extended to the multi-variate case. It is also possible to apply least-square inversion method to estimate the bootstrap samples (Efron, 1977). One thing to be considered in the multi-parameter least-square inversion (linear or non-linear) is the dependence between the parameters. If the parameters are linearly independent, the best estimates can be selected independently based on the analysis of each parameter by bootstrap sampling. On the other hand, if there is a trade-offs between parameters, the constraints between the parameters supersede the bootstrap analysis. In this case, the best estimate of each parameter is not necessarily coincided one another. The best estimates will be selected from an appropriate peak which satisfies an a priori information while maintaining the constraints.

Bootstrap with Non-linear Inversion

The property of the bootstrap sampling with outliers is especially important when the method is connected to the least-square inversion. It is notorious that the least-square inversion is not robust if there exist some outliers in the data. If the outliers in the data are either critical to the estimation of parameters or one-sided, the least-square inversion with bootstrap can give an idea of the median values of the data. On the other hand, the trade-off between parameters can cause an elongated distribution of the bootstrap estimation of parameters. For example in the source parameter inversion, the trade-off between steady state RDP and corner frequency may cause the elongation of steady state RDP distribution in one direction and that of the corner frequency in another direction. Therefore, it is dangerous to determine the optimal parameters by the information obtained from the bootstrap distribution. Optimal parameters should be chosen from the a priori information and the bootstrap estimation. The safest way to select an optimal parameters is to take the global maximum unless an a priori information does not exist and the distribution of the bootstrap estimation of parameters does not show clear peaks.

For the test of the bootstrap with non-linear inversion for the source parameter estimation, synthetic data were

generated from the Brune's model. Source parameters used to generate synthetic data were $\Psi_{\infty}=1$ and corner frequency=1.5 Hz. Attenuation($Q=30$) and random noise was applied to the synthetic data(Figure 3.68). Figure 3.68 also shows the result of single inversion as a thick line. The estimated parameters are steady state RDP=1.977, corner frequency=1.003 and $Q=33.45$ in a single inversion those of which are biased seriously by the noise. When inversion was taken 500 times from 500 bootstrap samples, the distribution of each parameter shows several distinctive peaks(Figure 3.69.a, b and c). Figure 3.70 is the plot of steady state RDP versus corner frequency, or trade-off map. An x-axis denotes steady state RDP and an y-axis denotes corner frequency. Each pair of steady state RDP and corner frequency is plotted as a dot. The plot shows the constraint between steady state RDP and corner frequency(slope). The slope by the trade-off between steady state RDP and corner frequency show slightly steeper value(-0.57) than the theoretical value(-0.5) but acceptable. The contour plot(Figure 3.71) shows maximum likelihood estimates of the parameters under constraint. There are three distinctive peaks each of which represents the best estimates of the sampling group, that is the group of relatively outlier-free sampling group, outlier-dominant group and the mean. The maximum likelihood estimates(global peak) corresponds to the mean. The estimates from the outlier-free sampling group which correspond to the median of

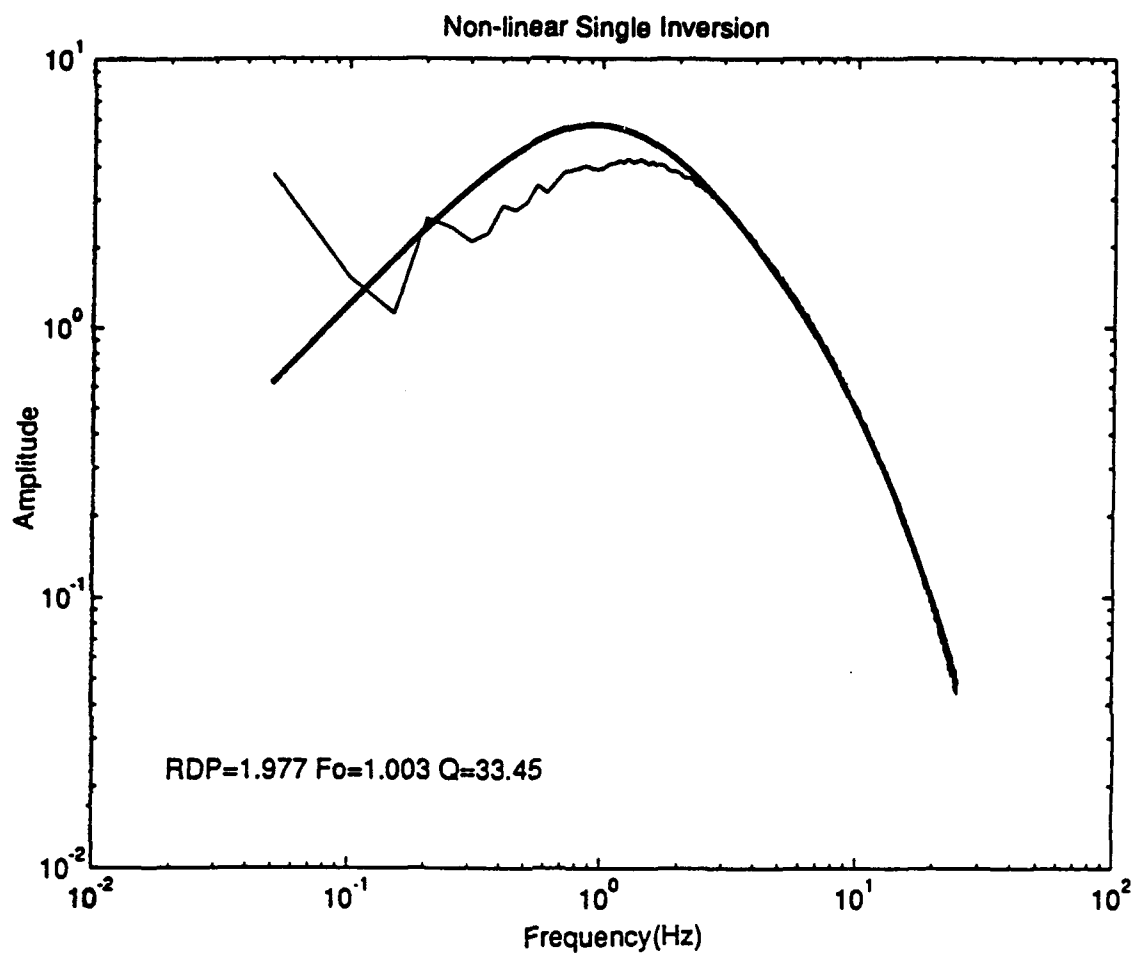
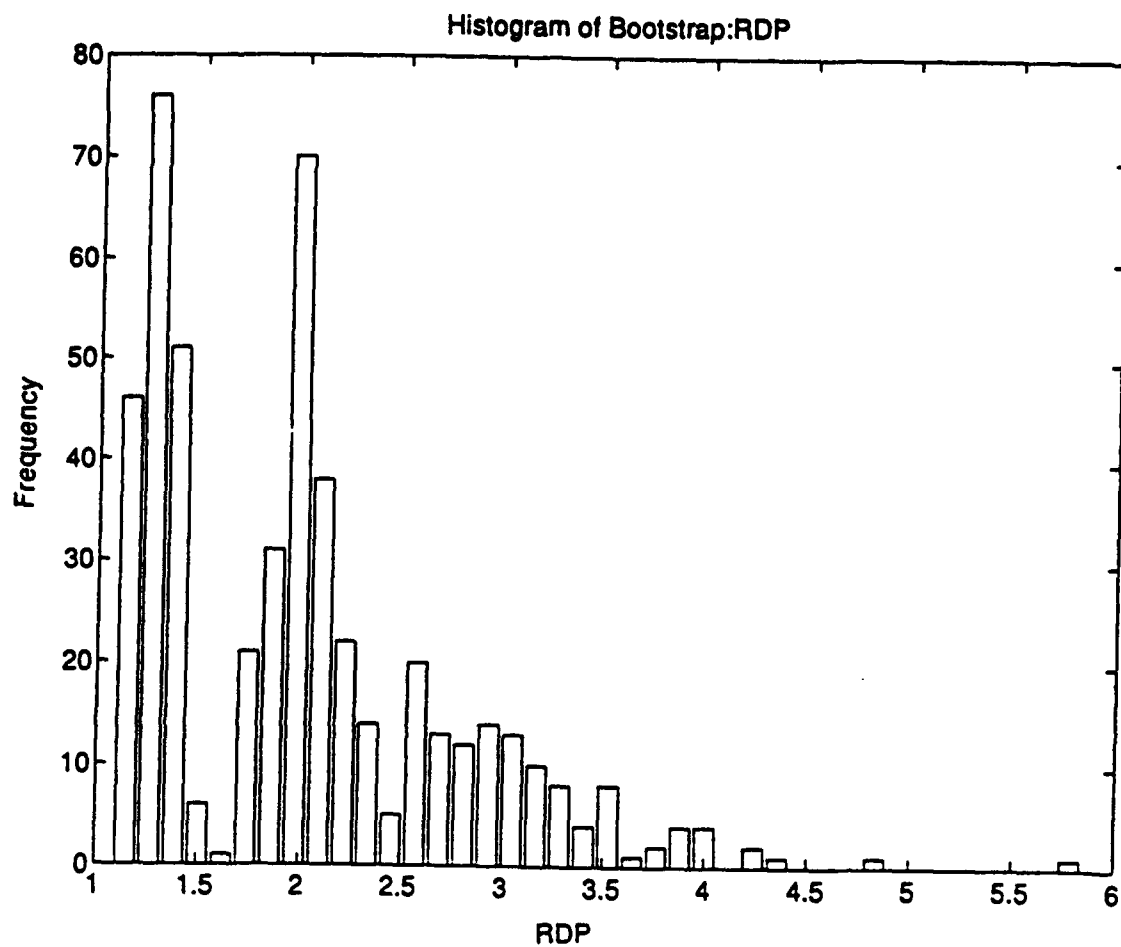
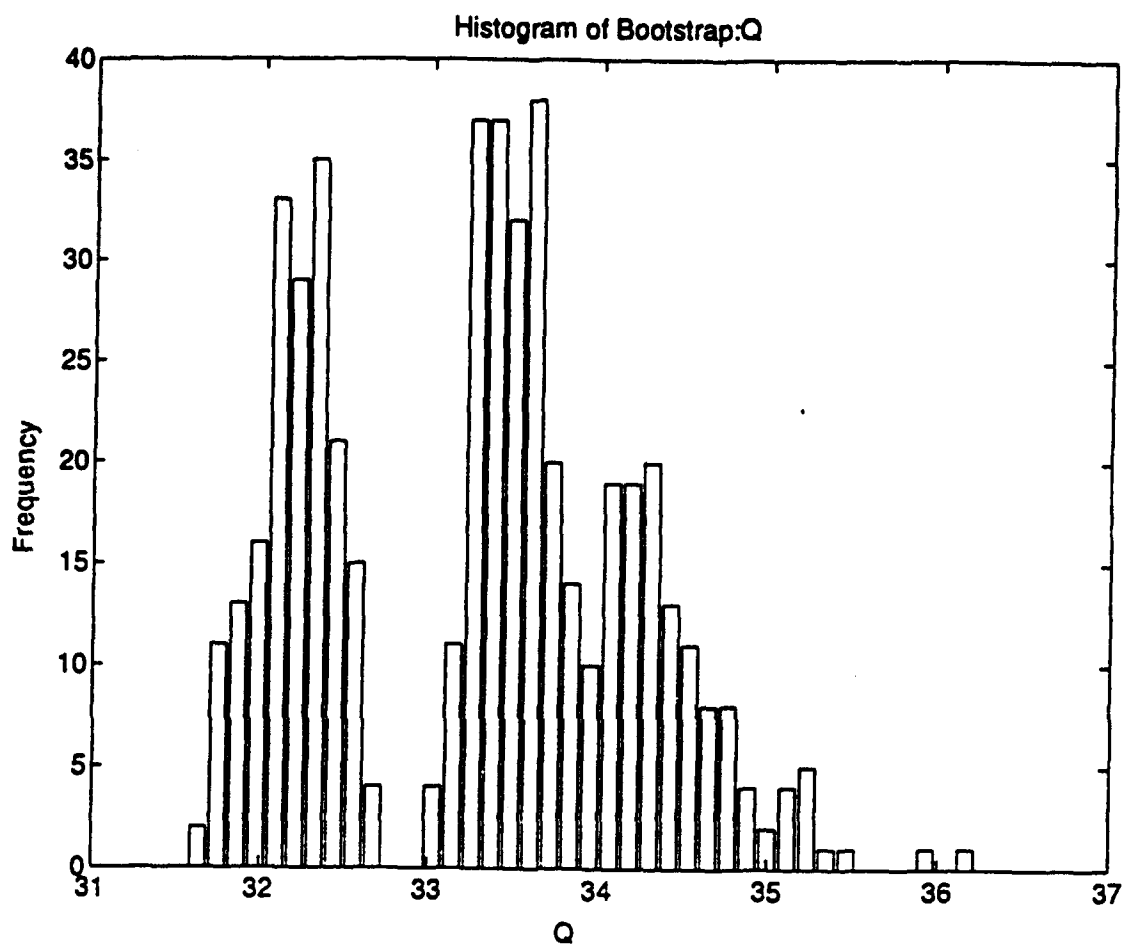


Figure 3.68 Result of a single non-linear inversion with the data with noise. Expected values are $RDP=1$, $F_o=1.5$, and $Q=30$.



(a)

Figure 3.69 Distribution of parameters by non-linear inversion with bootstrap sampling. 500 times of bootstrap sampling was done. (a)RDP distribution; (b)corner frequency distribution; (c)Q distribution.



(c)

Figure 3.69 Continued.

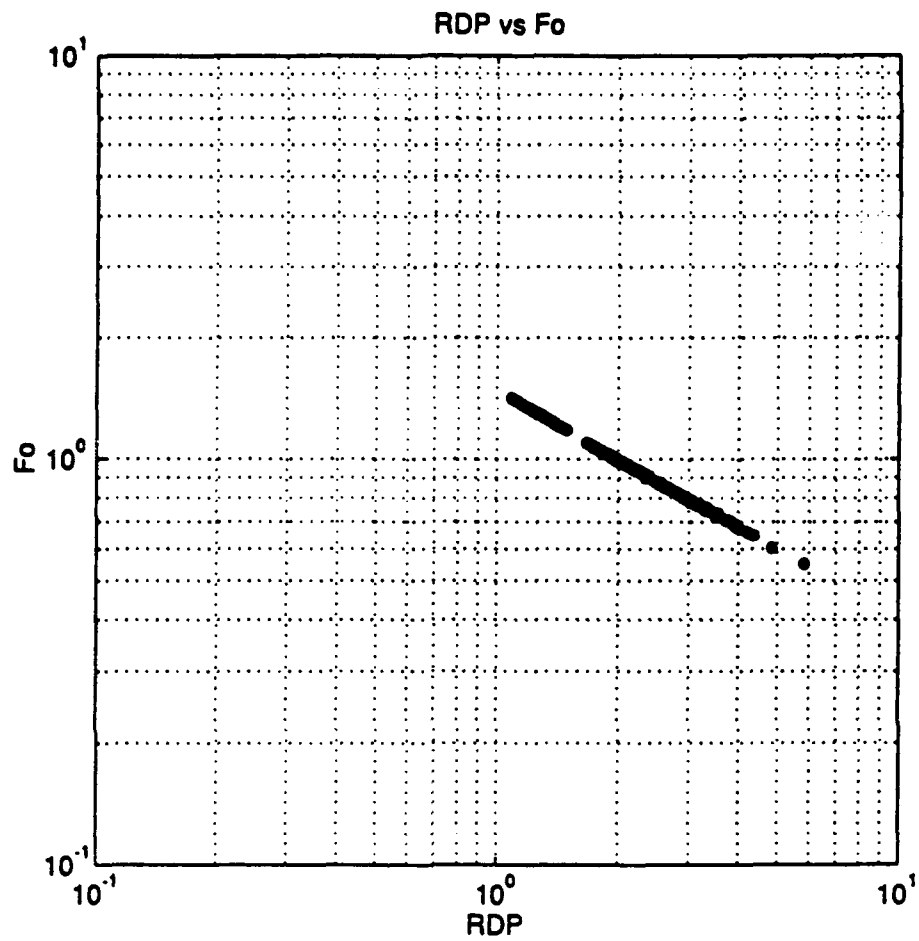


Figure 3.70 Trade-off map between RDP and corner frequency.

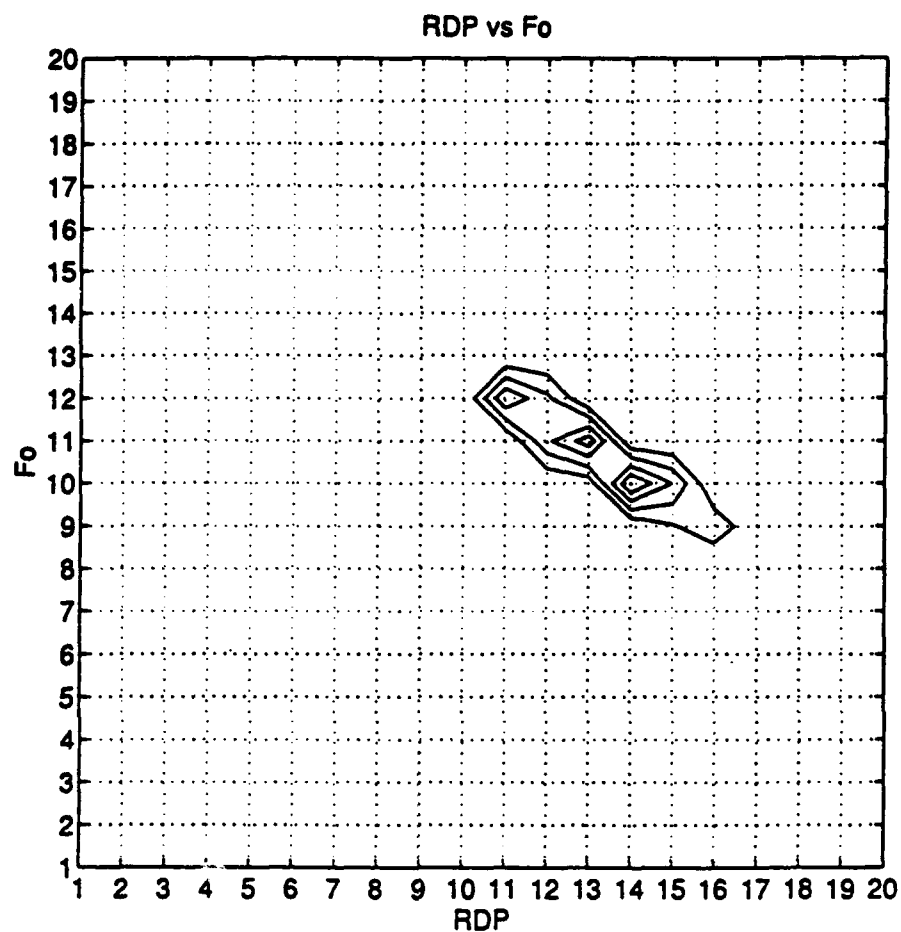


Figure 3.71 Contour map of Figure 3.73. There are three distinctive peaks within a given constraint.

the parameters are steady state RDP=1.2, corner frequency=1.3Hz and corresponding $Q=32.2$ each of which shows more robust and less biased result than that of parameter estimation with single non-linear inversion.

Synthetic seismograms generated by the reflectivity method with large attenuation($Q_0=30$) at 0.1 km and 1.0 km were used to test the applicability of the bootstrap, too. As a source function, von Seggern-Blanford type were used. The parameters used for the generation of source time function are $B=0$, $F_0=1.5$ Hz and $\Psi_\infty=0.01$ m³. Homogeneous full-space model was used as a path correction model. Figure 3.72 shows the synthetic data in the time domain. Single non-linear inversion estimate the parameters as $\Psi_\infty=0.0459$ and 0.0407, $F_0=0.8826$ and 0.5438, and $Q=9.8026$ and 20.65 respectively(Figure 3.73 a and b). These estimates are quite different from the expected values mainly because of the incorrect path effect compensation by the homogeneous full-space. 200 bootstrap estimates of the data at 0.1 km shows almost Gaussian distribution in its parameters(Figure 3.74 a, b, and c). The trade-off plot between steady state RDP and corner frequency shows a single peak(Figure 3.75 a and b). If we choose the global maximum as estimated parameters, steady state RDP is 0.045, corner frequency is 0.9 Hz and Q is 10.75 from the histogram of the Q . The bootstrap test for the data at 1.0 km shows the distribution with long tails in its steady state RDP and corner frequency estimation(Figure

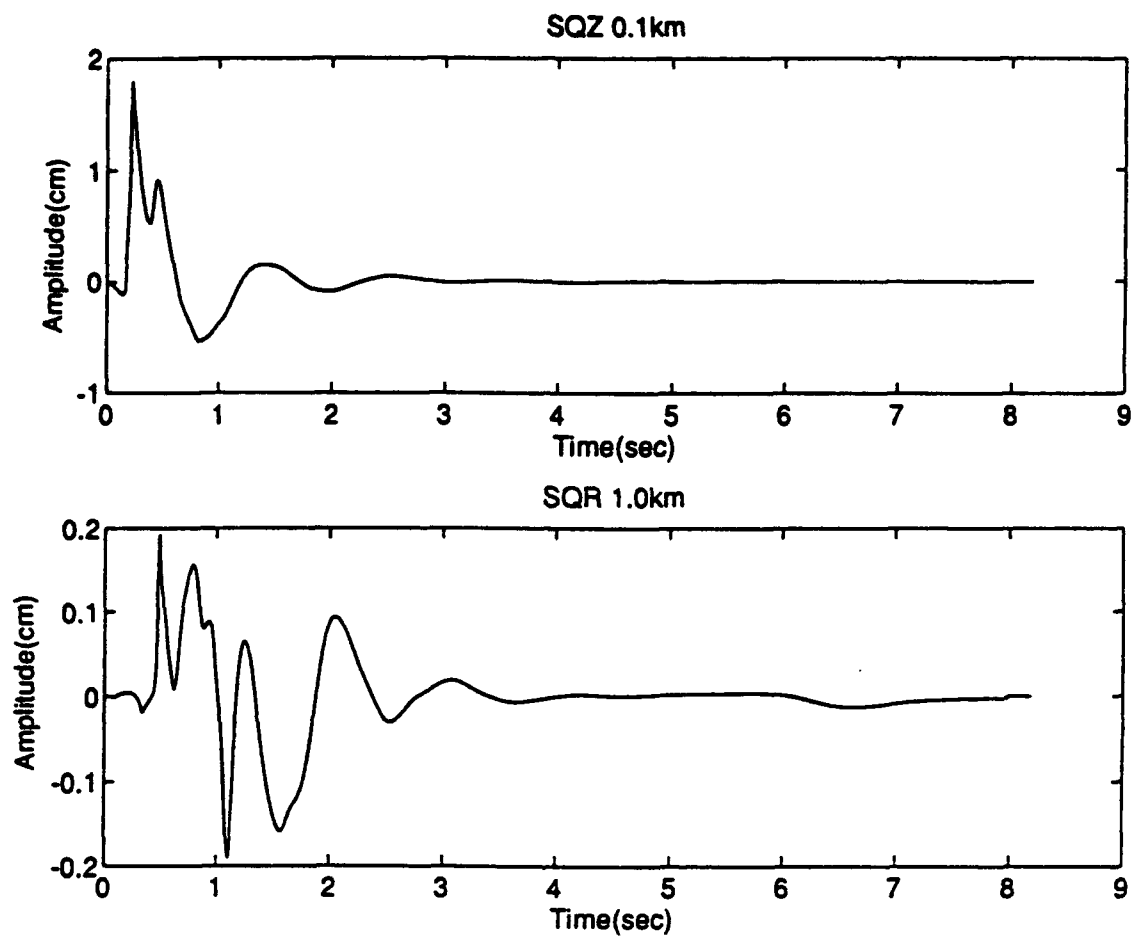


Figure 3.72 Synthetic seismograms used for the empirical test for bootstrap method.

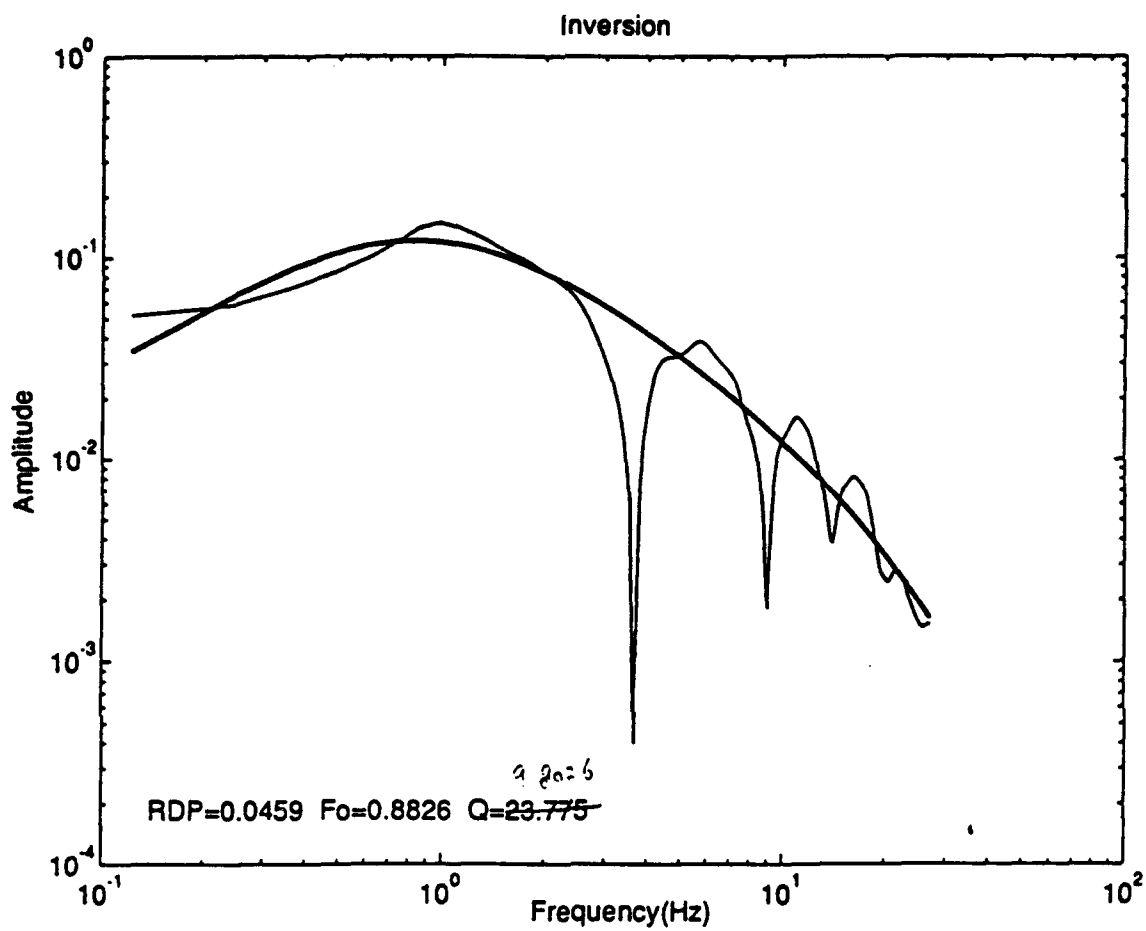
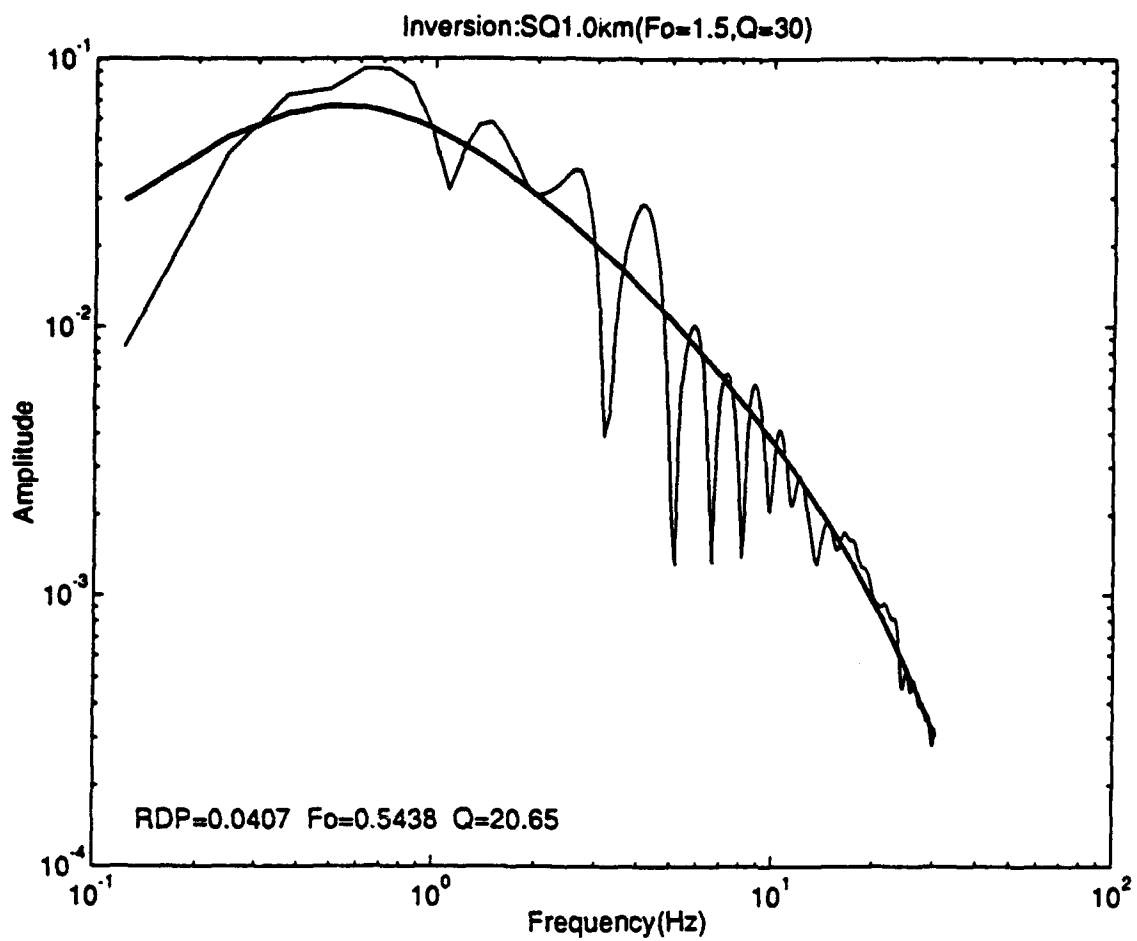
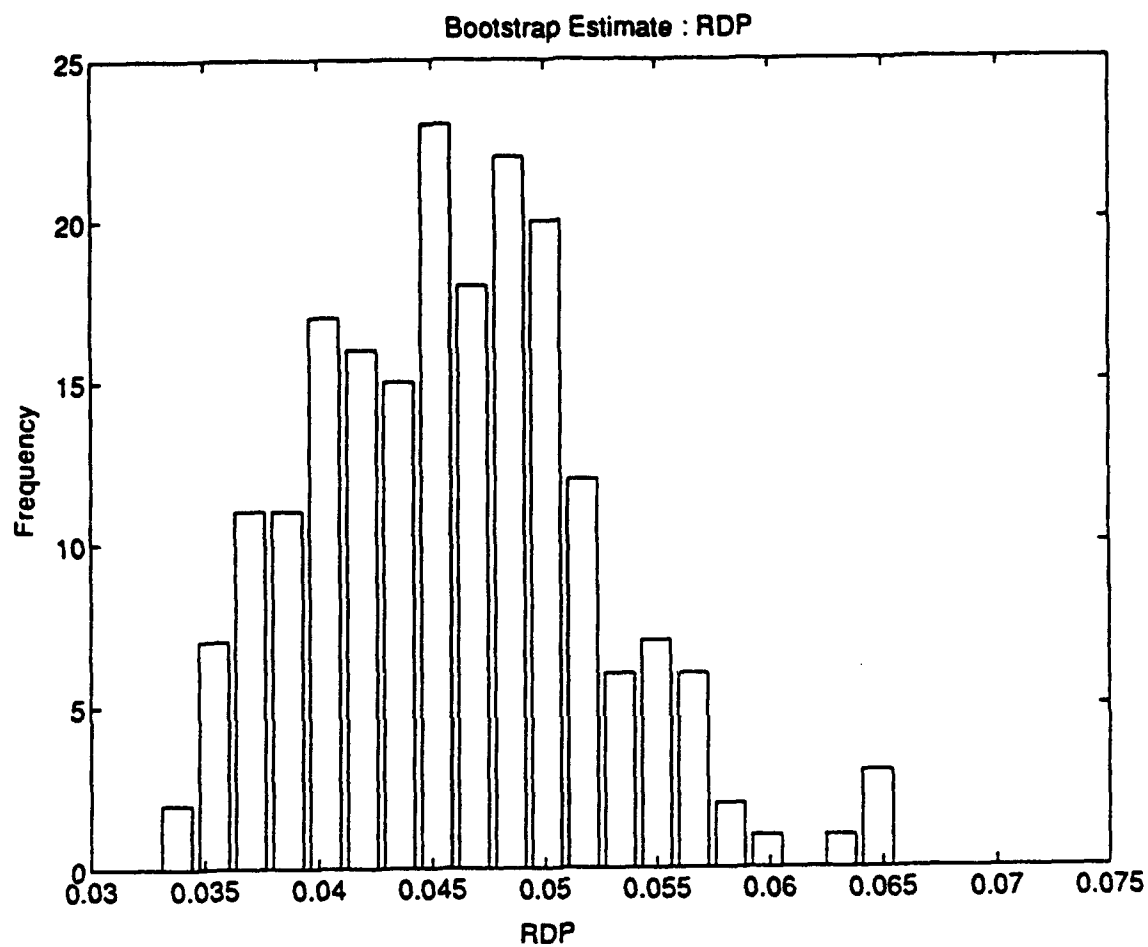


Figure 3.73 Single non-linear inversions from the synthetic seismograms. (a) SQZ 0.1 km; (b) SQR 1.0 km.



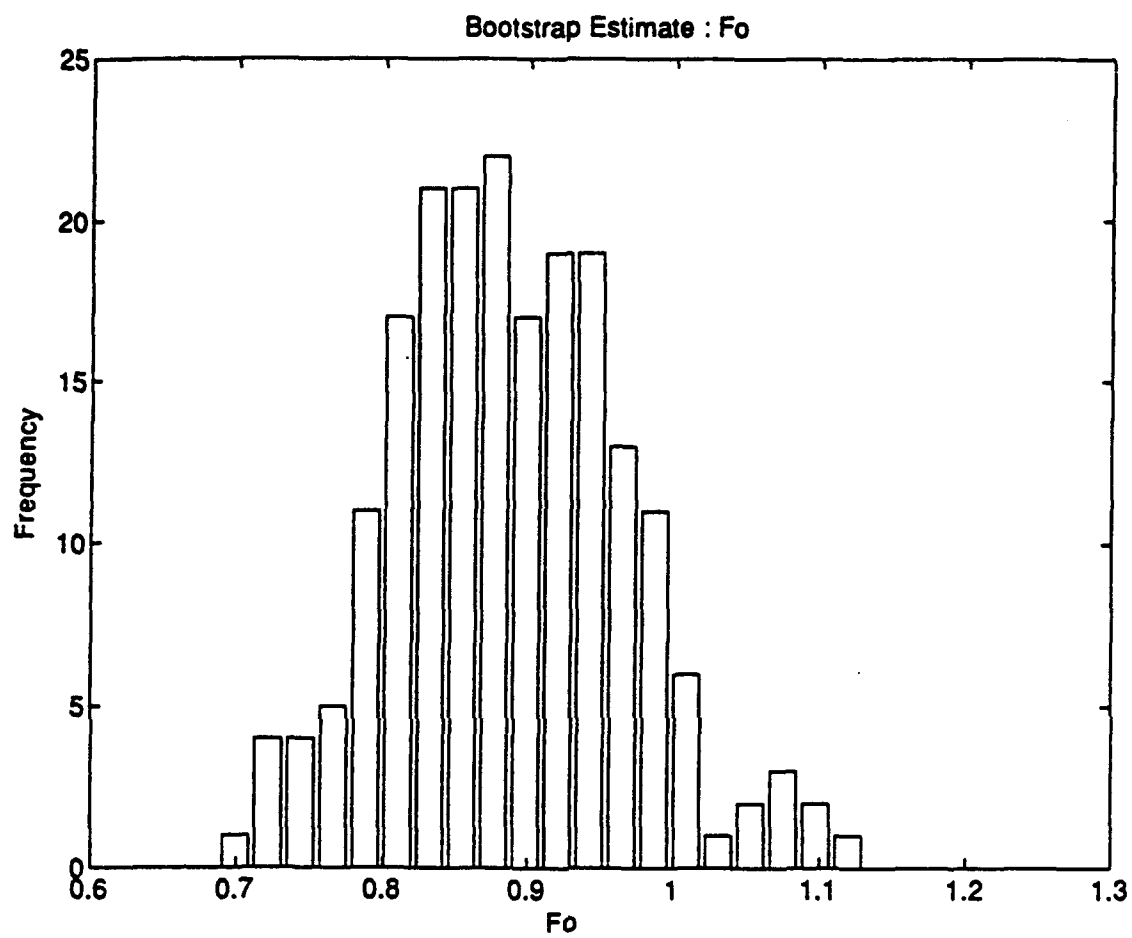
(b)

Figure 3.73 Continued.



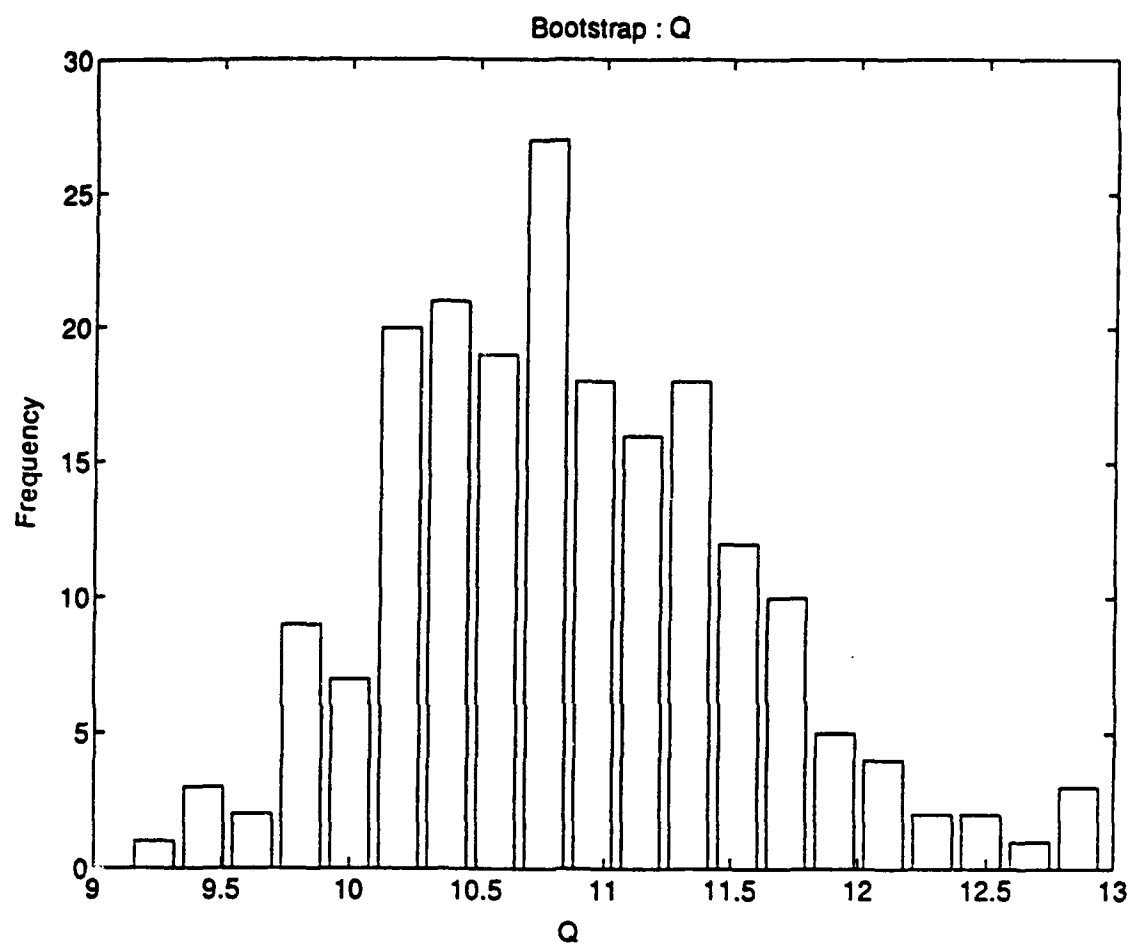
(a)

Figure 3.74 Distribution of parameters by non-linear inversion with bootstrap sampling. 200 bootstrap samples from SQZ(0.1 km) were used. (a)RDP distribution; (b)corner frequency distribution; (c)Q distribution.



(b)

Figure 3.74 Continued.



(c)

Figure 3.74 Continued.

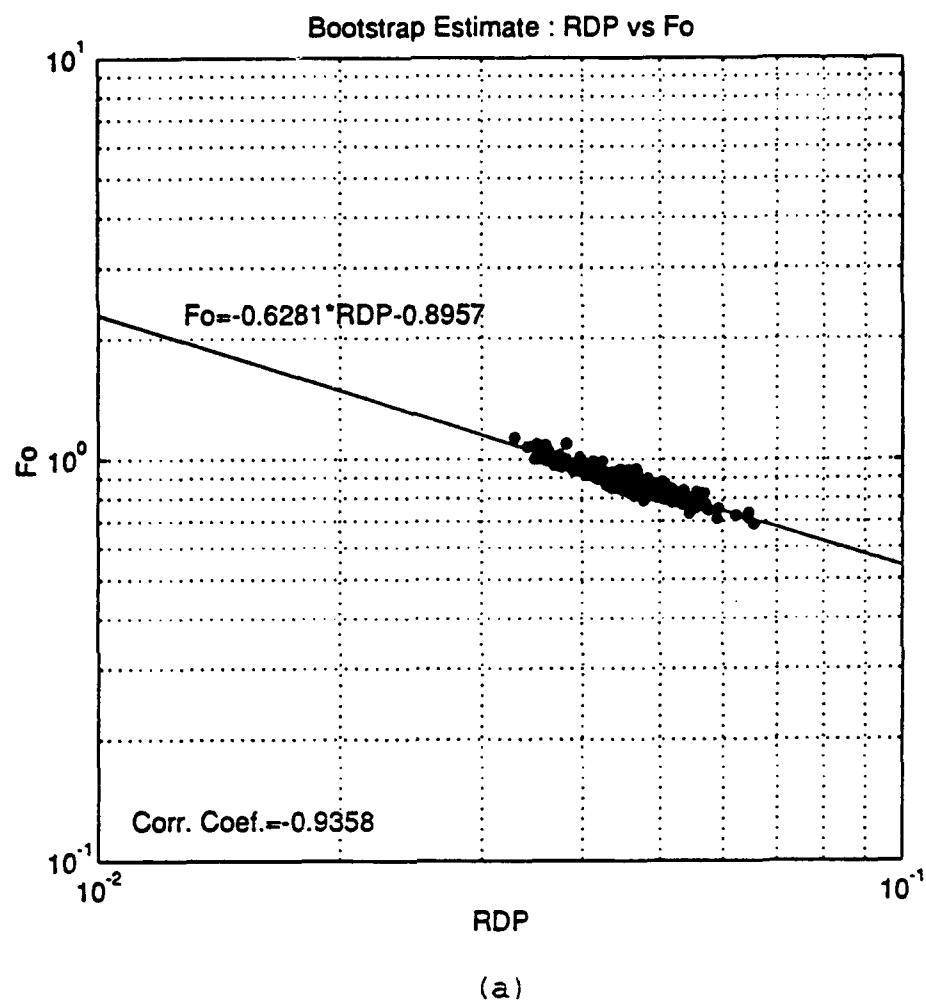
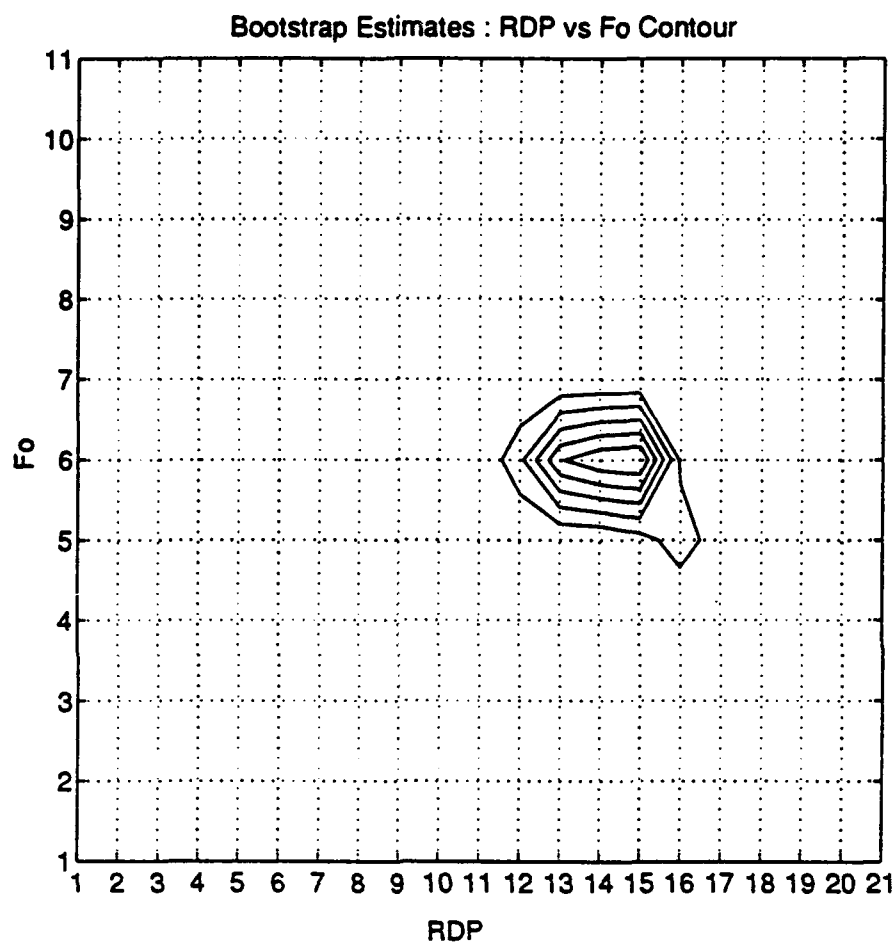


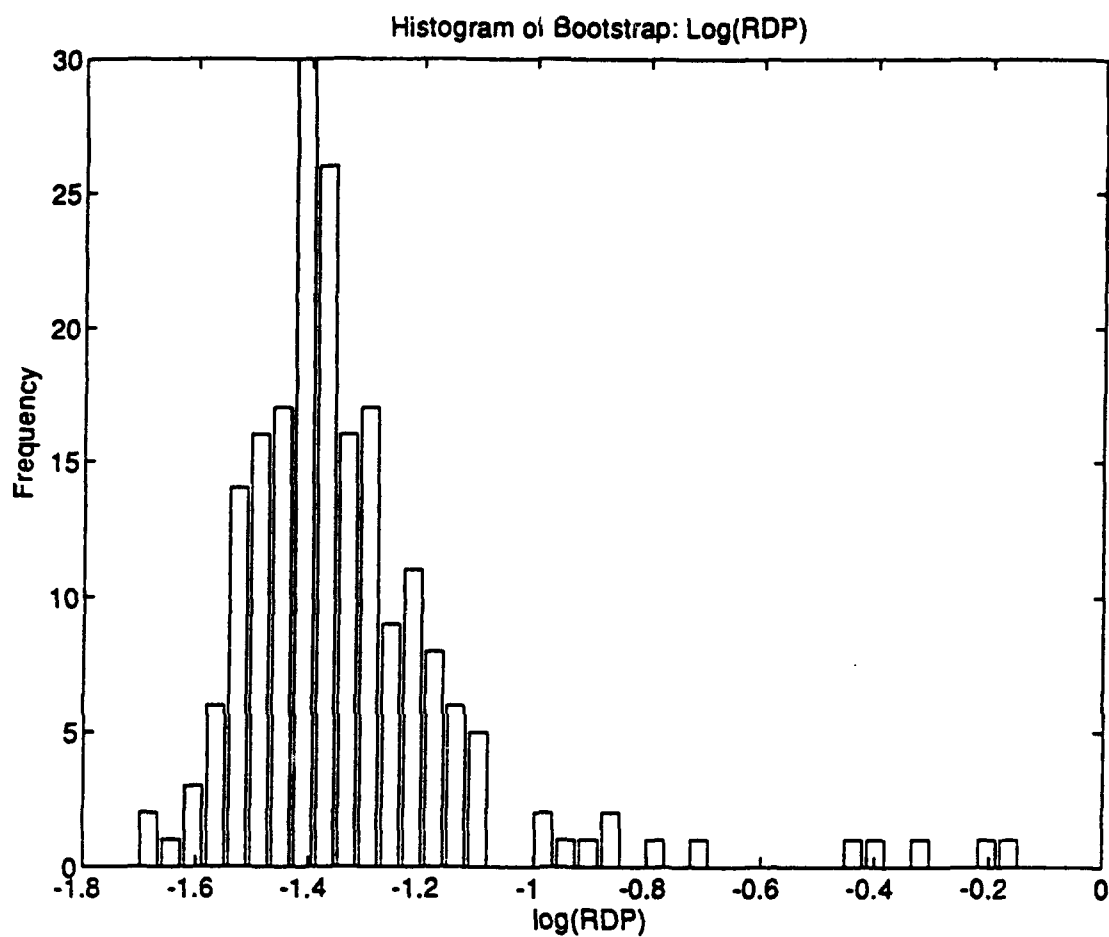
Figure 3.75 (a) Trade-off map between RDP and corner frequency. (b) Contour map. There is a distinctive peak within a given constraint.



(b)

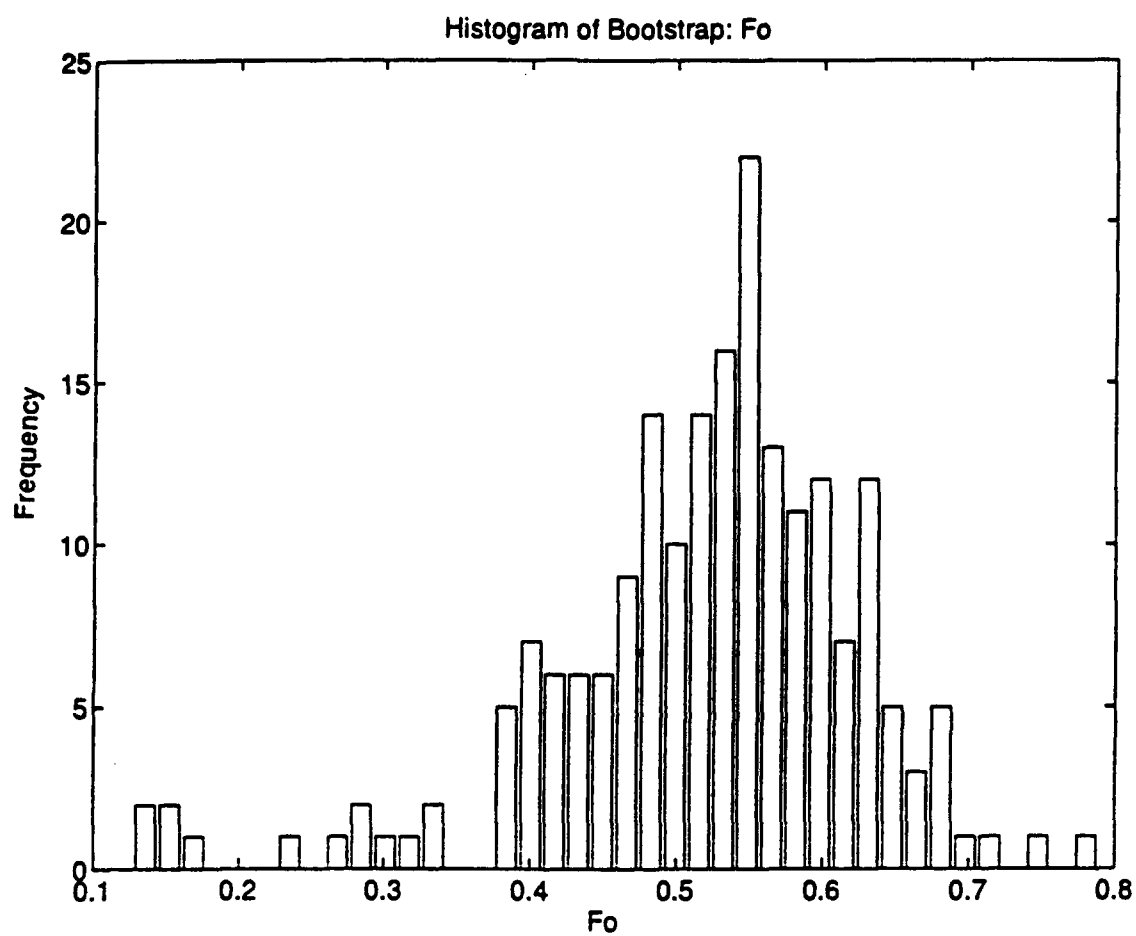
Figure 3.75 Continued.

3.76 a,b, and c) and three distinctive peaks in the contour plot(Figure 3.77 a and b). Except the longer tail which is resulted from the trade-off between parameters, the shape of the distribution is almost normal. If we choose the global maximum as an optimal estimates from the contour plot, the values are steady state RDP=0.04, f_o =0.57 and Q =20.8 respectively. It is not surprising that both data show Gaussian-like distribution since both are noise-free. When the noise is added in the data at 0.1 km(Figure 3.78), the distribution of bootstrap estimates shows different pattern(Figure 3.79 a, b, and c). Steady state and corner frequency distributions are not only far from Gaussian distribution but also headed in the same direction in their heads. The same direction of the heads in their skewed distribution implies that the skewness of the distribution is not related to the trade-off. It is also possible to select two peaks in the corner frequency distribution. On the other hand the distribution of the Q shows Gaussian distribution. The trade-off curves(Figure 3.80 a and b) also show three peaks which are quite different in its pattern from the data without noise(Figure 3.75 a and b). If the peak which shows smallest steady state RDP and largest corner frequency in the contour plot is picked as an optimal estimates, based on the distribution of each parameter, the values are steady state RDP=0.08, f_o =0.63, and Q =12.5 which are better estimation than the single non-linear inversion(steady state RDP=0.14,



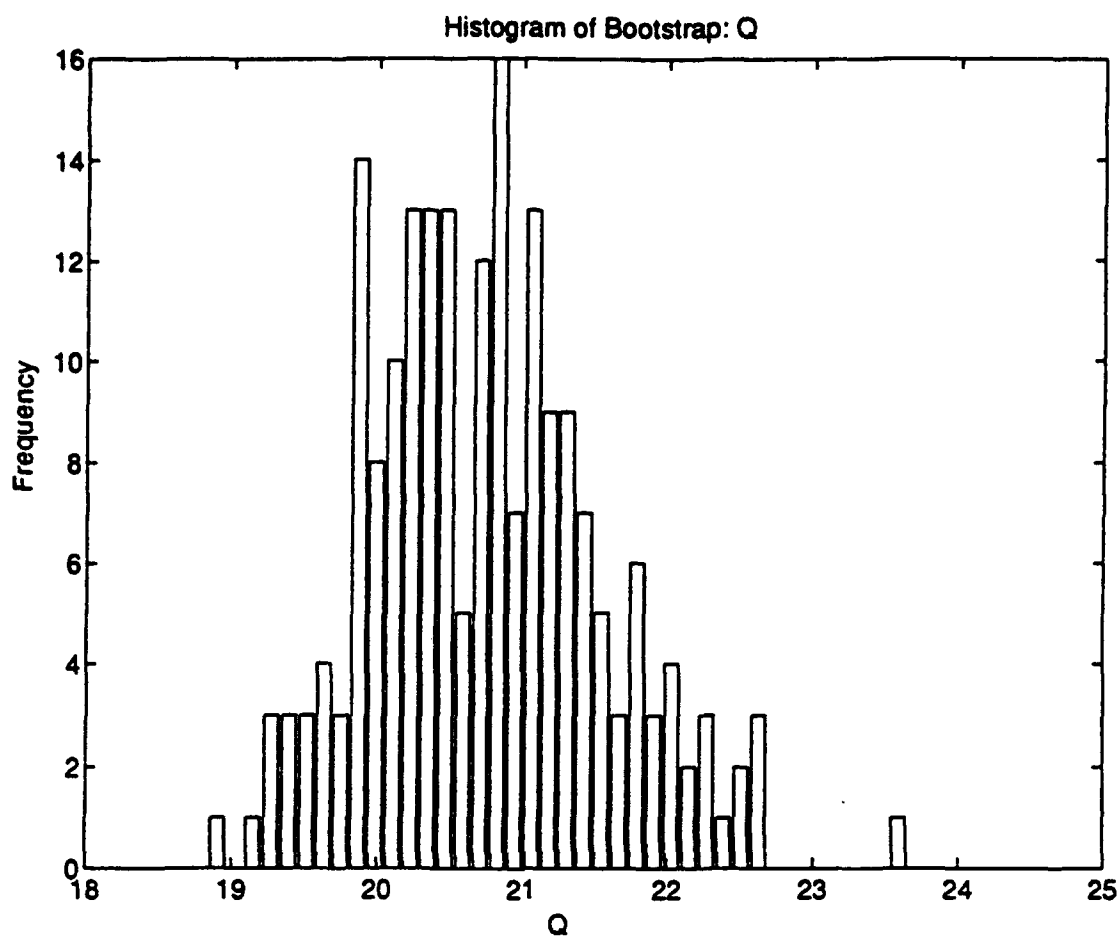
(a)

Figure 3.76 Distribution of parameters by non-linear inversion with bootstrap sampling. 200 bootstrap samples from SQR(1.0 km) were used. (a)RDP distribution; (b)corner frequency distribution; (c)Q distribution.



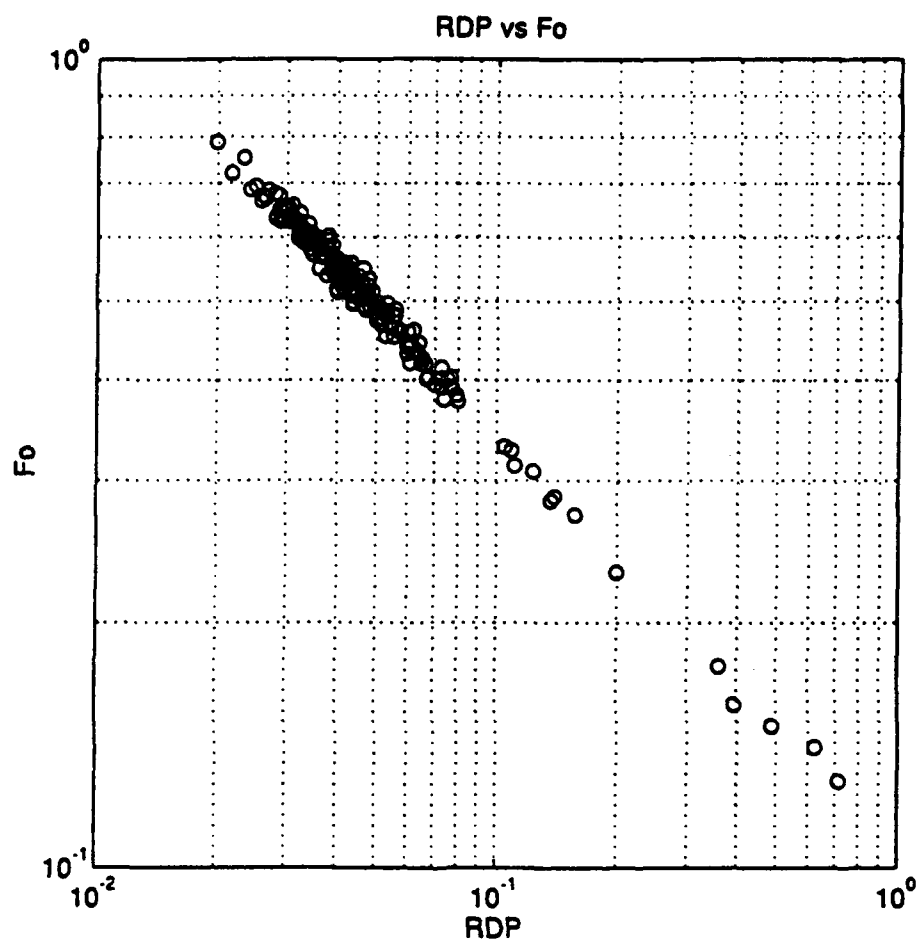
(b)

Figure 3.76 Continued.



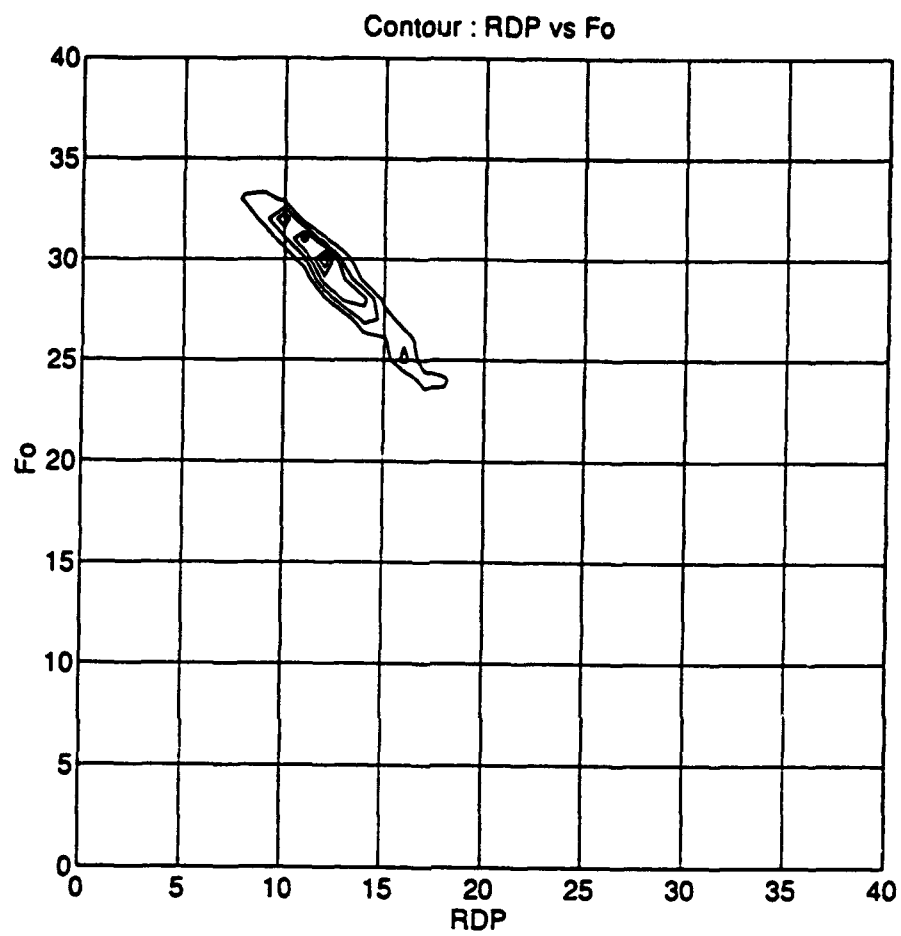
(c)

Figure 3.76 Continued.



(a)

Figure 3.77 (a) Trade-off map between RDP and corner frequency. (b) Contour map. There are distinctive peaks within a given constraint.



(b)

Figure 3.77 Continued.

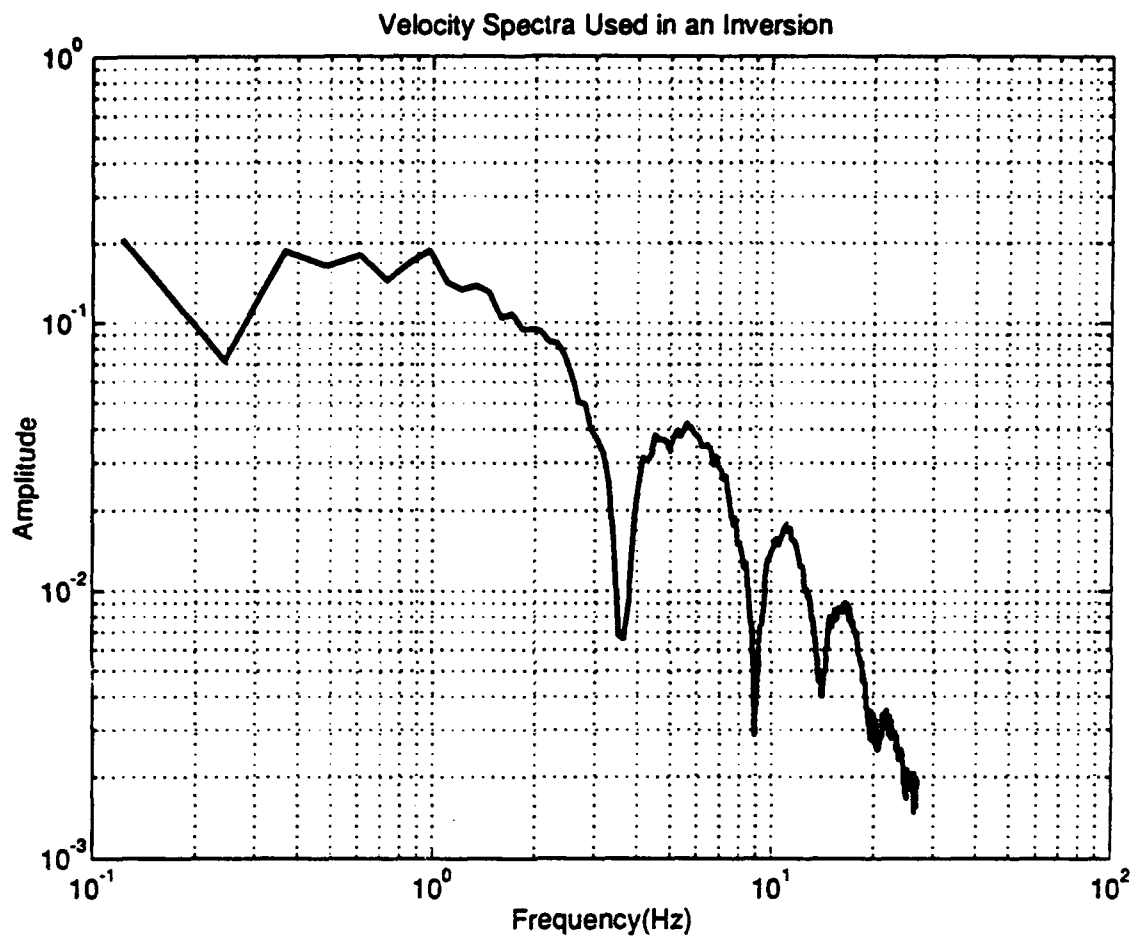
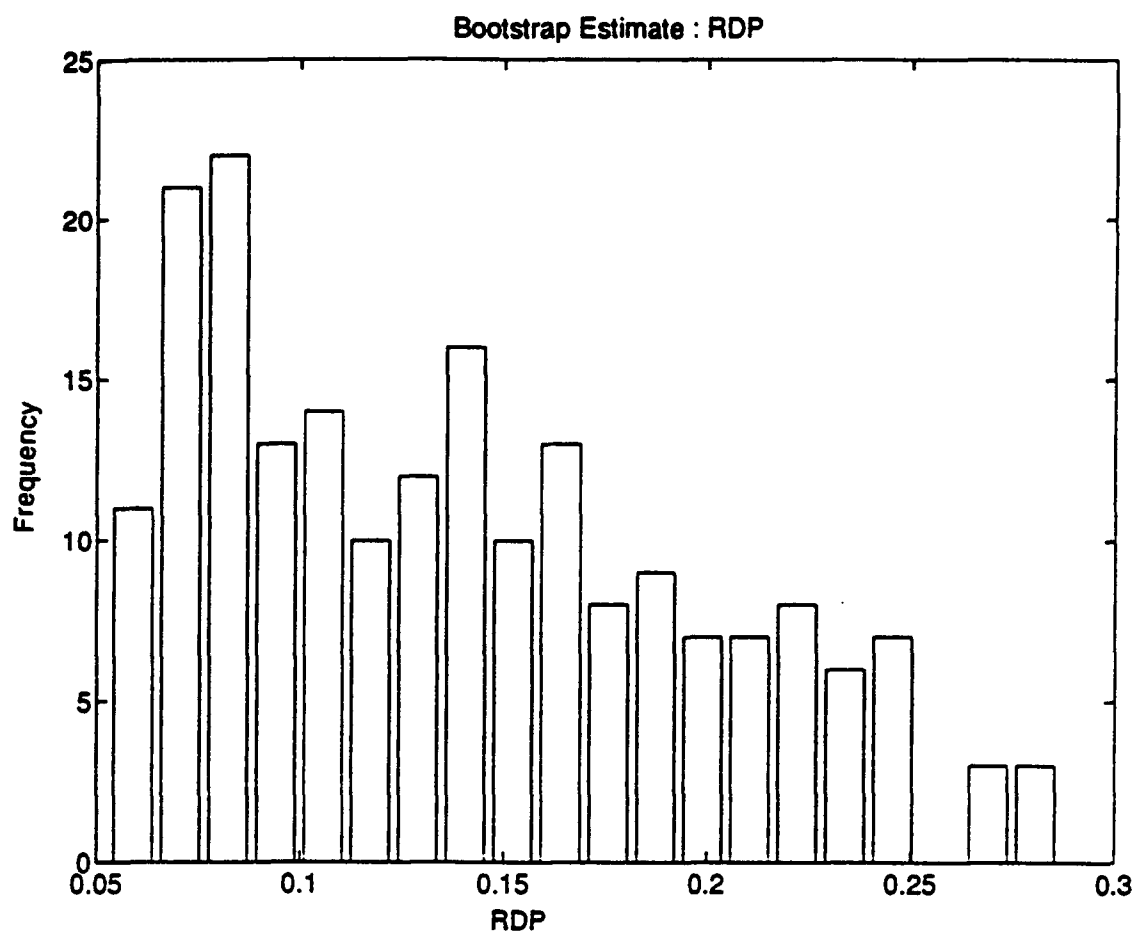
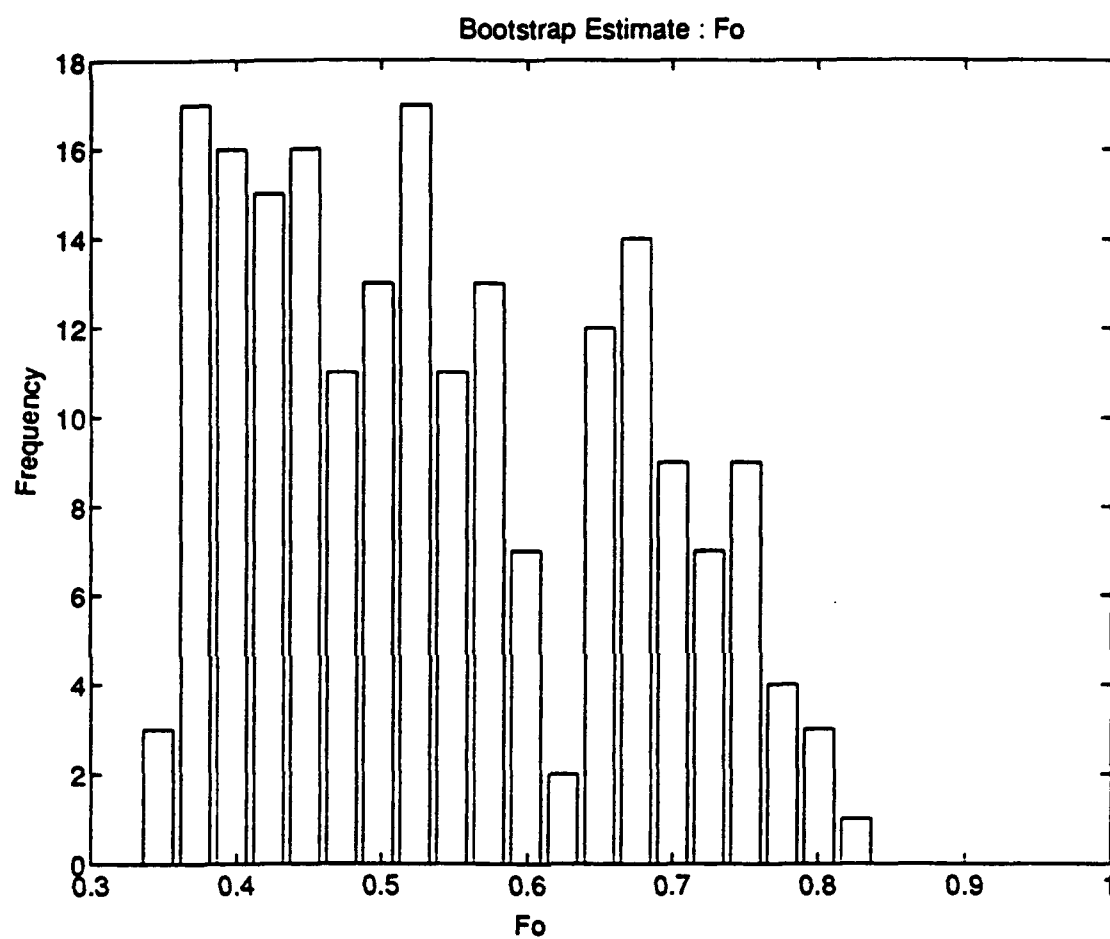


Figure 3.78 Spectra from synthetic seismograms (SQZ 0.1 km) with random noise. Random noise has the slope of -1 which biases the low- and high-frequencies.



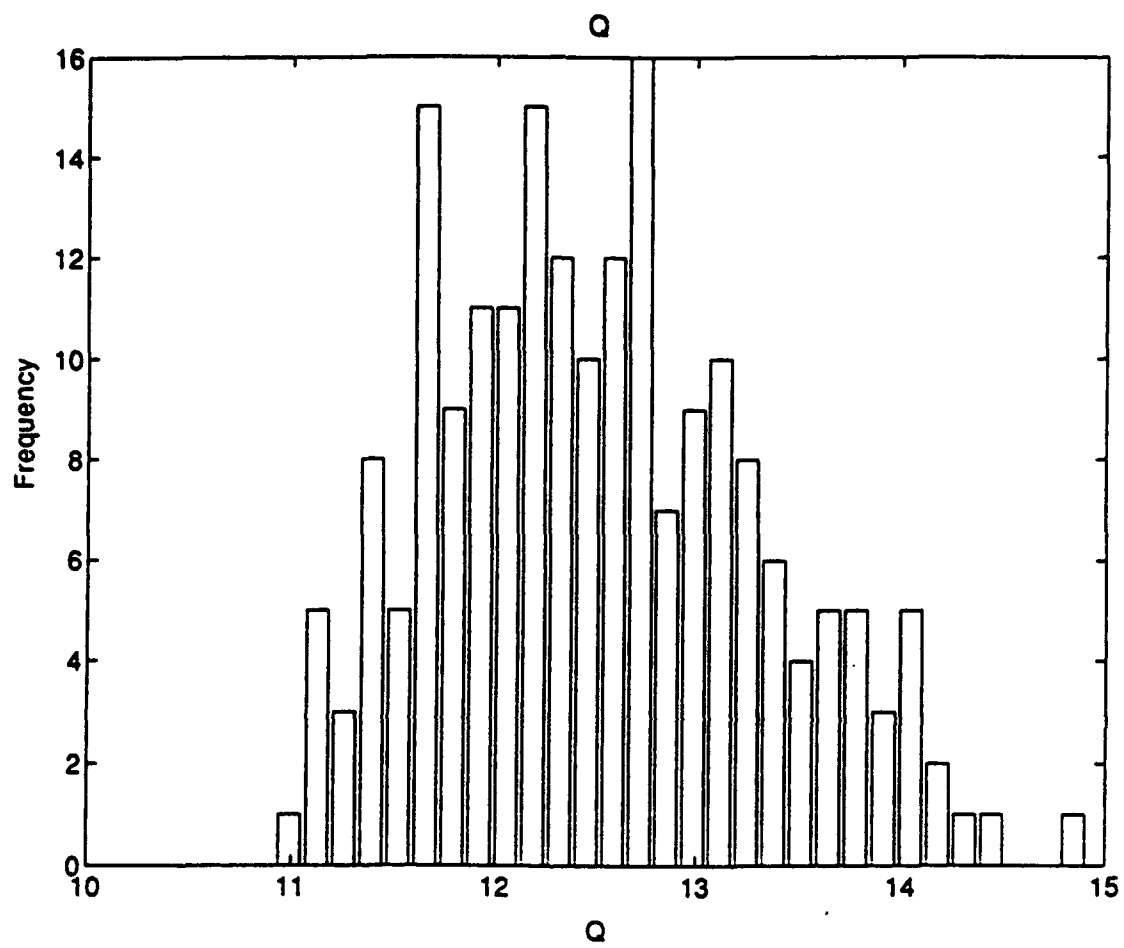
(a)

Figure 3.79 Distribution of parameters by non-linear inversion with bootstrap sampling. 200 bootstrap samples from noise contaminated SQZ(0.1 km) were used. (a)RDP distribution; (b)corner frequency distribution; (c)Q distribution.



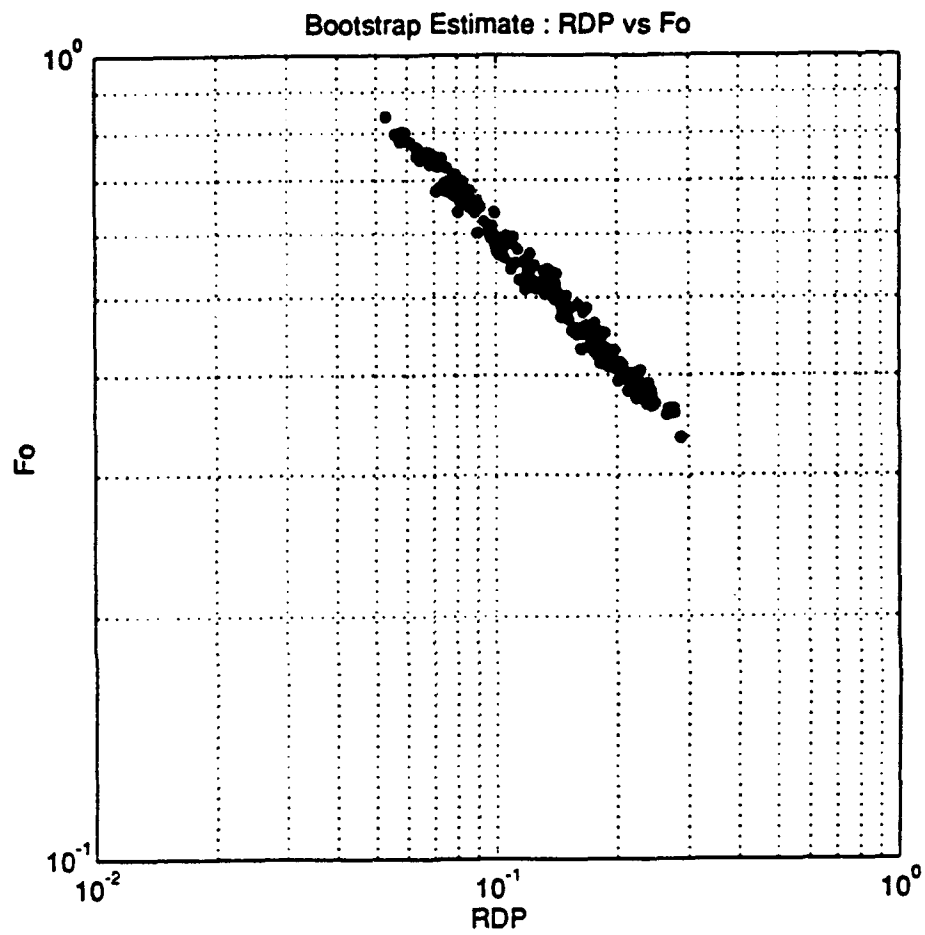
(b)

Figure 3.7c Continued.



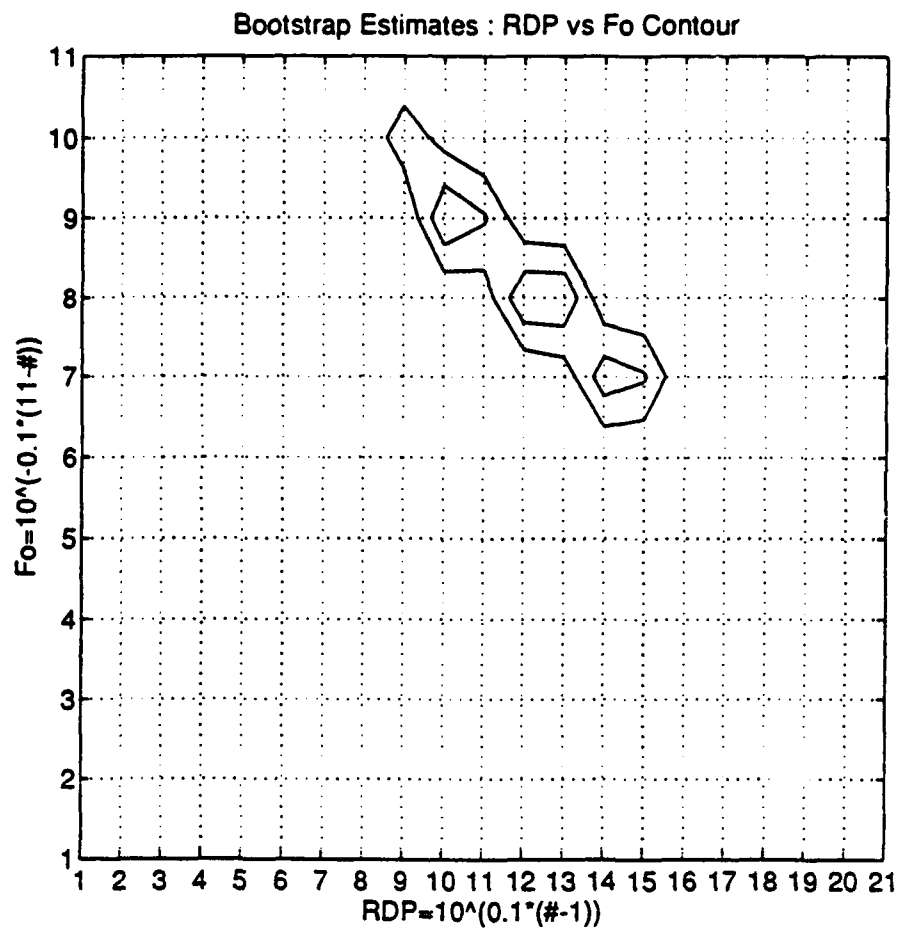
(c)

Figure 3.79 Continued.



(a)

Figure 3.80 (a) Trade-off map between RDP and corner frequency. (b) Contour map. There are three distinctive peaks within a given constraint.



(b)

Figure 3. $\S 0$ Continued.

$f_0=0.50$, $Q=11.4$). Noise added for the test was uniformly distributed random noise which had a slope of -1 with respect to the frequency.

Summary

The inversion schemes introduced in the previous chapter were tested with synthetically generated data. Each scheme has its own merits and demerits. Based on the various posterior analysis such as resolution and covariance matrices, condition number and degree of bias, the scheme with prewhitening and damping (Scheme 4) was proved to be optimal to investigate the source parameters. This scheme was modified to the scheme with varying damping parameter (Modified scheme) for the speed of inversion. The empirical tests illustrate that the modified scheme works impeccably for the source parameter inversion.

Empirical tests and the posterior analysis show that the von Seggern-Blanford model is less stable than the Brune's model due to the ill-posedness of overshoot. A small amount of error introduced at low frequencies results in a large variation in the VSB model's STEADY STATE RDP and overshoot. On the other hand, the forward model without overshoot (Brune's model), shows consistent estimates with some bias.

Path effects such as near-field and surface waves effect were tested with synthetic data generated from the

homogeneous full-space, homogeneous half-space and layered half-space. Path correction was done by homogeneous full-space, homogeneous half-space and layer over infinite half-space models. This test illustrate that the source parameters can be biased seriously by the path effect. Near-field term strongly biases STEADY STATE RDP within a range of a few wavelengths. Surface wave effect is so devastating that it scrambles every parameter in the forward source model. The tests by the Helmberger-Hadley's model for the data generated with the von Seggern-Blanford's model illustrate that it is impossible to prefer any source model without correcting the path effects, especially surface wave effect.

High frequency range approximated by STEADY STATE $RDP \cdot (1+2B) \cdot F_0^2$ in the von Seggern-Blanford model or by STEADY STATE $RDP \cdot F_0^2$ in the Brune's model shows consistent estimates throughout the whole data set. This frequency band is not affected seriously by the near-field, spall, noise and surface wave effects.

There are two advantages in using bootstrap method for the estimation of source parameters using non-linear inversion. The first is related to the proven characteristic of bootstrap. Since some of the source parameters are determined from a limited number of data point, the robustness of the bootstrap method in estimating statistics from a limited data can help to estimate them reliably.

Another advantage of bootstrap is, as shown in the empirical tests, that this method indicates the reliability of the estimated parameters. Asymmetric ratio, quantitative measurement of skewness, is a good indicator of the reliability as shown in the empirical tests. If the outliers bias the statistics of the sample, the statistics of bootstrap sample show either multiple peak or skewed distribution toward the maximum outlier from synthetically generated random number test. When the synthetic data are used for the source parameter inversion with bootstrapping, the bootstrap estimation shows characteristic distribution according to the existence of the noise. Less reliable data set from the bootstrap estimation, which may cause large variance in the resolution of source parameters, can be excluded based on the skewness of the bootstrap estimates. The distribution of the bootstrap estimates may indicate the degree of bias and the optimal estimates under certain circumstances. One-dimensional random number test shows that there is a peak corresponding to the median of the samples, which is believed to be a better estimate if small number of outliers exist. This characteristic, however, needs more theoretical investigation and is not applied to estimate source parameters in this paper.